

NON-PROVISIONAL

**PATENT APPLICATION**

**PAPER-BASED INTERFACE FOR MULTIMEDIA INFORMATION  
STORED BY MULTIPLE MULTIMEDIA DOCUMENTS**

Inventor(s): Jonathan J. Hull, a citizen of The United States, residing at  
751 Laurel Street PMB 434  
San Carlos, CA 94070

Jamey Graham, a citizen of The United States, residing at  
1196 Shasta Avenue  
San Jose, CA 95126

Assignee: Ricoh Company, Ltd.  
15-5 Minami Aoyama 1-Chome  
Minato-Ku  
Tokyo, JAPAN 107-8544

Entity: Large

**PAPER-BASED INTERFACE FOR MULTIMEDIA INFORMATION  
STORED BY MULTIPLE MULTIMEDIA DOCUMENTS**

**COPYRIGHT**

[01] A portion of the disclosure of this patent document contains material  
5 that is subject to copyright protection. The copyright owner has no objection to the  
xerographic reproduction by anyone of the patent document or the patent disclosure in  
exactly the form it appears in the U.S. Patent and Trademark Office patent file or records, but  
otherwise reserves all copyright rights whatsoever.

**CROSS-REFERENCES TO RELATED APPLICATIONS**

[02] The present application incorporates by reference for all purposes the  
entire contents of the following applications:

[03] (1) U.S. Non-Provisional Patent Application No. 08/995,616 filed  
December 22, 1997;

[04] (2) U.S. Non-Provisional Patent Application No. 09/149,921 filed  
September 9, 1998;

[05] (3) U.S. Non-Provisional Patent Application No. \_\_/\_\_,\_\_,  
(Attorney Docket No. 15358-006500US) entitled "PAPER-BASED INTERFACE FOR  
MULTIMEDIA INFORMATION" filed concurrently with this application;

20 [06] (4) U.S. Non-Provisional Patent Application No. \_\_/\_\_,\_\_,  
(Attorney Docket No. 15358-007200US) entitled "TECHNIQUES FOR RETRIEVING  
MULTIMEDIA INFORMATION USING A PAPER-BASED INTERFACE" filed  
concurrently with this application;

[07] (5) U.S. Non-Provisional Patent Application No. \_\_/\_\_,\_\_,  
25 (Attorney Docket No. 15358-007300US) entitled "TECHNIQUES FOR ANNOTATING  
MULTIMEDIA INFORMATION" filed concurrently with this application; and

[08] (6) U.S. Non-Provisional Patent Application No. \_\_/\_\_,\_\_,  
(Attorney Docket No. 15358-007500US) entitled "TECHNIQUES FOR GENERATING A  
COVERSHEET FOR A PAPER-BASED INTERFACE FOR MULTIMEDIA  
30 INFORMATION" filed concurrently with this application.

## BACKGROUND OF THE INVENTION

[09] The present invention relates to techniques for accessing multimedia information, and more particularly to techniques for generating a printable representation of the multimedia information that can be printed on a paper medium to provide a paper-based interface for the multimedia information.

[10] With the rapid growth of computers, an increasing amount of information is being stored in the form of electronic (or digital) documents. These electronic documents include multimedia documents that store multimedia information. The term "multimedia information" is used to refer to information that comprises information of several different types in an integrated form. The different types of information included in multimedia information may include a combination of text information, graphics information, animation information, sound (audio) information, video information, and the like. Multimedia information is also used to refer to information comprising one or more objects wherein the objects include information of different types. For example, multimedia objects included in multimedia information may comprise text information, graphics information, animation information, sound (audio) information, video information, and the like.

[11] Several different techniques and tools are available today for accessing and navigating multimedia information that may be stored in electronic multimedia documents. Examples of such tools and/or techniques include proprietary or customized multimedia players (e.g., RealPlayer™ provided by RealNetworks, Microsoft Windows Media Player provided by Microsoft Corporation, QuickTime™ Player provided by Apple Corporation, Shockwave multimedia player, and others), video players, televisions, personal digital assistants (PDAs), and the like.

[12] The tools and techniques described above that are conventionally available for accessing multimedia information focus on the electronic or digital nature/format of the multimedia information and output the multimedia information in electronic or digital form. For example, multimedia players typically execute on a computer system and output the multimedia information stored in multimedia documents via output devices coupled to the computer such as a monitor, a speaker, and the like.

[13] While retrieving multimedia information in digital form is adequate for some users, it is a well-known fact that many users find it easier to comprehend and assimilate information when the information is printed on a paper medium rather than in the digital form. These users thus prefer to access information in a paper format by printing the

information on a paper medium. For example, most people who encounter a long document will typically print the document on paper before reading the document, even though there are several tools (e.g., word processors, browsers, etc.) available for viewing and navigating the document in electronic form. While there are several tools available for printing ordinary data files containing text and images on paper (e.g., a printer coupled to a word-processor), there are no techniques or tools that allow users to print multimedia information on a paper-medium in a format and style that is readable by the user. As described above, all of the conventionally available tools and techniques for retrieving multimedia information focus on the electronic or digital nature/format of the multimedia content and output the multimedia information in electronic or digital form.

[14] In light of the above, there is a need for techniques that allow users to access multimedia information via a paper-based interface.

#### BRIEF SUMMARY OF THE INVENTION

[15] The present invention provides techniques for generating a printable representation of multimedia information that can be printed on a paper medium to provide a paper-based interface for the multimedia information. According to an embodiment of the present invention, the present invention provides techniques for generating a single printable representation that includes multimedia information extracted from a plurality of different multimedia documents or multimedia sources. The single printable representation includes multimedia information selected from the plurality of multimedia documents based upon selection criteria. The selection criteria may be based upon any attributes of the multimedia documents or their contents, or upon user-specified topics of interest, and the like. The single or consolidated printable representation can then be printed on a paper medium to generate a consolidated paper document comprising information that satisfies the selection criteria.

[16] According to an embodiment of the present invention, technique are provided for generating a paper document based upon a plurality of multimedia documents storing multimedia information in electronic form. The embodiment of the present invention receives input identifying a selection criterion. The embodiment analyzes the multimedia information stored by the plurality of multimedia documents to identify portions of multimedia information that satisfy the selection criterion, the portions of multimedia information including at least a first portion extracted from a first multimedia document from the plurality of multimedia documents and a second portion extracted from a second multimedia document from the plurality of multimedia documents. The portions of the



multimedia information that satisfy the selection criterion are then printed on a paper medium to generate the paper document comprising a set of one or more printed pages.

[17] According to another embodiment of the present invention, techniques are provided for generating a paper document using multimedia information stored by a first multimedia document and a second multimedia document. The embodiment of the present invention receives input identifying a selection criterion and accesses printable representations for the first multimedia document and the second multimedia document. The present invention analyzes the printable representation for the first multimedia document to identify at least one portion of the printable representation that satisfies the selection criterion, and analyzes the printable representation for the second multimedia document to identify at least one portion of the printable representation that satisfies the selection criterion. A consolidated printable representation is then generated that includes the at least one portion of the printable representation for the first multimedia document and the at least one portion for the second multimedia document that satisfy the selection criterion. The consolidated printable representation is then printed on a paper medium to generate the paper document comprising one or more printed pages.

[18] According to another embodiment of the present invention, a paper document is provided that comprises one or more pages, wherein at least one page of the one or more pages is imprinted with text information that is extracted from multimedia information stored by a plurality of multimedia documents if the text information satisfies a selection criterion, and wherein the at least one page is imprinted with one or more video frames corresponding to the text information extracted from the plurality of multimedia documents.

[19] The foregoing, together with other features, embodiments, and advantages of the present invention, will become more apparent when referring to the following specification, claims, and accompanying drawings.

#### BRIEF DESCRIPTION OF THE DRAWINGS

[20] Fig. 1 is a simplified block diagram of a distributed system that incorporates an embodiment of the present invention;

[21] Fig. 2A depicts a networked system including a multifunction device according to an embodiment of the present invention;

[22] Fig. 2B depicts a user interface that is displayed to the user by a multifunction device according to an embodiment of the present invention;

[23] Fig. 3 is a simplified block diagram of a computer system according to an embodiment of the present invention.;

[24] Fig. 4 is a simplified high-level flowchart depicting a method of generating a printable representation of multimedia information according to an embodiment of the present invention;

[25] Figs. 5A and 5B depict a sample template according to an embodiment of the present invention;

[26] Fig. 6 is a simplified high-level flowchart depicting processing performed in step 408 of Fig. 4 according to an embodiment of the present invention;

[27] Fig. 7A depicts a page from a multimedia paper generated according to an embodiment of the present invention for a multimedia document;

[28] Fig. 7B depicts a second page that follows the page depicted in Fig. 7A in a multimedia paper document according to an embodiment of the present invention;

[29] Fig. 7C depicts a page from a multimedia paper generated according to an embodiment of the present invention for a multimedia document;

[30] Fig. 8A, 8B, and 8C depict pages from a multimedia paper document generated for a recorded meeting according to an embodiment of the present invention;

[31] Figs. 9A, 9B, and 9C depict pages of a multimedia paper document displaying visual markers to denote various attributes of the audio information or of the CC text information included in the multimedia information for the multimedia document for which the multimedia paper document is generated according to an embodiment of the present invention;

[32] Fig. 10 depicts a page from a multimedia paper document whose contents have been annotated according to an embodiment of the present invention;

[33] Fig. 11 depicts a user profile that may be configured by a user according to an embodiment of the present invention to facilitate selection of keyframes relevant to user-specified topics of interest;

[34] Fig. 12 depicts modules that facilitate selection of keyframes relevant to topics of interest according to an embodiment of the present invention;

[35] Fig. 13A is a simplified high-level flowchart depicting a method of accessing multimedia information using a multimedia paper document according to an embodiment of the present invention;

[36] Fig. 13B is a simplified high-level flowchart depicting a method of accessing multimedia information from a particular time point using a multimedia paper document according to an embodiment of the present invention;

5 [37] Fig. 14 is a simplified high-level flowchart depicting a method of generating a single printable representation according to an embodiment of the present invention that includes multimedia information selected from a plurality of multimedia documents by analyzing the printable representations of the plurality of multimedia documents;

[38] Fig. 15 is a simplified high-level flowchart depicting another method of generating a single printable representation that includes information extracted from a plurality of multimedia documents by analyzing the multimedia information stored by the plurality of multimedia documents according to an embodiment of the present invention;

[39] Figs. 16A, 16B, 16C, and 16D depict pages of a multimedia paper document generated according an embodiment of the present invention using the method depicted in Fig. 14;

[40] Fig. 17 depicts a coversheet generated for a multimedia paper document according to an embodiment of the present invention;

[41] Fig. 18 depicts a coversheet generated for a multimedia paper document according to another embodiment of the present invention;

[42] Fig. 19 depicts a coversheet generated according to another embodiment of the present invention for a multimedia paper document that has been annotated based upon user-specified topics of interest;

25 [43] Fig. 20 depicts a coversheet generated according to an embodiment of the present invention for a multimedia paper document that includes pages selected from multiple multimedia paper documents based upon selection criteria;

[44] Fig. 21 depicts another coversheet generated according to an embodiment of the present invention for a multimedia paper document that includes pages selected from multiple multimedia paper documents based upon selection criteria; and

30 [45] Fig. 22 depicts a coversheet generated according to an embodiment of the present invention for a multimedia paper document that has been generated for a recorded meeting.

## DETAILED DESCRIPTION OF THE INVENTION

[46] The present invention provides techniques for generating a printable representation of multimedia information that can be printed on a paper medium to provide a paper-based interface for the multimedia information. The paper-based interface provided by the present invention provides a more readable and comprehensible representation of the multimedia information.

[47] According to an embodiment of the present invention, the printable representation for the multimedia information may be annotated to identify locations of information in the multimedia information that may be of interest to a user. A paper document generated by printing the annotated printable representation on a paper medium displays the annotations. The annotations provide visual indications of information relevant to the user. For example, information printed in the paper document that is relevant to topics of interest specified by a user may be annotated or highlighted. In this manner, the multimedia paper document generated according to the teachings of the present invention provides a convenient tool that allows a user to readily locate portions of the paper document that are relevant to the user. Since the multimedia paper document comprises a printable representation of multimedia information, the paper document generated according to the teachings of the present invention allows the user to identify portions of multimedia information that are of interest to the user.

[48] According to an embodiment of the present invention, the paper document generated by printing the printable representation on a paper medium also provides an interface for accessing or retrieving multimedia information in electronic form. The paper document may thus be used as an indexing and retrieval tool for retrieving multimedia information. For example, a user may use a paper document generated for a video recording to access or retrieve portions of the video recording.

[49] According to an embodiment of the present invention, the present invention provides techniques for generating a single printable representation that includes multimedia information extracted from a plurality of different multimedia documents or multimedia sources. According to an embodiment of the present invention, the single printable representation includes multimedia information selected from the plurality of multimedia documents based upon selection criteria. A user may specify the selection criteria. The selection criteria may be based upon any attributes of the multimedia documents or their contents, or upon user-specified topics of interest, and the like. The single or

consolidated printable representation can then be printed on a paper medium to generate a consolidated paper document comprising information that satisfies the selection criteria.

[50] According to an embodiment of the present invention, the present invention provides techniques for generating a coversheet for a paper document generated by printing the printable representation on a paper medium. The coversheet may provide a summary of the contents printed on pages of the paper document.

[51] As described above, the printable representation of the multimedia information can be printed on a paper medium to generate the paper-based interface. The term "paper" or "paper medium" as used in this application is intended to refer to any tangible medium on which information can be printed, written, drawn, imprinted, embossed, etc. For purposes of this invention, the term "printing" is intended to include printing, writing, drawing, imprinting, embossing, and the like. For purposes of this invention, the document generated by printing the printable representation on a paper medium will be referred to as "multimedia paper" or "multimedia paper document." The multimedia paper document takes advantage of the high resolution and portability of paper and provides a readable representation of the multimedia information. According to the teachings of the present invention, a multimedia paper document may also be used to select, retrieve, and access the multimedia information.

[52] The multimedia information for which the multimedia paper document is generated may be stored in an electronic multimedia document. Accordingly, the term "multimedia document" is intended to refer to any storage unit (e.g., a file) that stores multimedia information in digital format. Various different formats may be used to store the multimedia information. These formats include various MPEG formats (e.g., MPEG 1, MPEG 2, MPEG 4, MPEG 7, etc.), MP3 format, SMIL format, HTML+TIME format, WMF (Windows Media Format), RM (Real Media) format, Quicktime format, Shockwave format, various streaming media formats, formats being developed by the engineering community, proprietary and customary formats, and others. Examples of multimedia documents include video recordings, MPEG files, news broadcast recordings, presentation recordings, recorded meetings, classroom lecture recordings, broadcast television programs, and the like.

[53] As previously described, multimedia information comprises information of different types in an integrated form. For example, multimedia information may comprise a combination of text, graphics, animation, sound (audio), and/or video information in an integrated form. For example, a video recording of a television broadcast may comprise video information and audio information. In certain instances the video

recording may also comprise close-captioned (CC) text information which comprises material related to the video information, and in many cases, is an exact representation of the speech contained in the audio portions of the video recording. As another example, a recording of a presentation may store information captured during a presentation including video

5 information, audio information, CC text information, information corresponding to slides presented during the presentation, whiteboard information, and other types of information. As described below, the present invention generates a printable representation of the multimedia information that includes printable representations of the various types of information included in the multimedia information. The printable representation of the multimedia document can then be printed on a paper medium to generate a multimedia paper or multimedia paper document for the multimedia information stored by the multimedia document.

#### [54] GENERATING PRINTABLE REPRESENTATION OF MULTIMEDIA INFORMATION

[55] As described above, according to an embodiment of the present invention, techniques are provided for generating a printable representation of multimedia information that can be printed on a paper medium to produce a multimedia paper document. The multimedia paper document provides a paper-based interface for the user to view and comprehend the multimedia information. Fig. 1 is a simplified block diagram of a distributed system 100 that might incorporate an embodiment of the present invention. As depicted in Fig. 1, distributed system 100 comprises a number of devices or computer systems including one or more user systems 102, a multimedia information processing server system (MIPSS) 104, a multimedia information source (MIS) 106, and a multimedia paper output device 108  
25 coupled to communication network 110 via a plurality of communication links. It should be apparent that distributed system 100 depicted in Fig. 1 is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives. For example, in alternative embodiments of the present invention, one or more of the systems depicted in Fig. 1 (e.g., MIPSS 104 and output device  
30 108) may be incorporated into a single system. In other alternative embodiments, the present invention may also be embodied in a stand-alone system, and the like.

[56] Communication network 110 provides a mechanism allowing the various devices and computer systems depicted in Fig. 1 to communicate and exchange data

and information with each other. Communication network 110 may itself be comprised of many interconnected computer systems and communication links. While in one embodiment, communication network 110 is the Internet, in other embodiments, communication network 110 may be any suitable communication network including a local area network (LAN), a wide area network (WAN), a wireless network, an intranet, a private network, a public network, a switched network, and the like.

[57] The communication links used to connect the various systems depicted in Fig. 1 may be of various types including hardwire links, optical links, satellite or other wireless communications links, wave propagation links, or any other mechanisms for communication of information. Various communication protocols may be used to facilitate communication of information via the communication links. These communication protocols may include TCP/IP, HTTP protocols, extensible markup language (XML), wireless application protocol (WAP), protocols under development by industry standard organizations, vendor-specific protocols, customized protocols, and others.

[58] According to the teachings of the present invention, MIPSS 104 is configured to perform processing to facilitate generation of a printable representation of the multimedia information. The printable representation generated by MIPSS 104 for a multimedia document may include printable representations of the various types of information included in the multimedia information stored by the multimedia document. The printable representation generated by MIPSS 104 may be printed on a paper medium to generate a multimedia paper document. The processing performed by MIPSS 104 to generate a printable representation of the multimedia information may be implemented by software modules executing on MIPSS 104, by hardware modules coupled to MIPSS 104, or combinations thereof. According to alternative embodiments of the present invention, the processing may also be distributed between other computer systems and devices depicted in Fig. 1.

[59] The multimedia information for which MIPSS 104 generates a printable representation may be stored in a multimedia document accessible to MIPSS 104. For example, the multimedia document may be stored by MIPSS 104 or may alternatively be stored in locations accessible to MIPSS 104.

[60] In alternative embodiments of the present invention, instead of being stored in a multimedia document, MIPSS 104 may receive a stream of multimedia information (e.g., a streaming media signal, a cable signal, etc.) from a multimedia information source such as MIS 106. Examples of MIS 106 include a television broadcast

receiver, a cable receiver, a TIVO box, and the like. MIPSS 104 may receive the multimedia information directly from MIS 106 or may alternatively receive the information via a communication network such as communication network 106. MIPSS 104 may then store the multimedia information received from MIS 106 in the form of a multimedia document and use the stored information to generate the printable representation of the multimedia information.

[61] After generating the printable representation of the multimedia information, MIPSS 104 may communicate the printable representation to output device 108 that is capable of generating a multimedia paper document by printing the printable representation on a paper medium. In one embodiment, MIPSS 104 may itself be configured to generate a multimedia paper document from the printable representation of the multimedia information. In alternative embodiments, the printable representation generated by MIPSS 104 may be stored for later use.

[62] As described above, multimedia information source (MIS) 106 represents a source of multimedia information. According to an embodiment of the present invention, MIS 106 may store multimedia documents that are accessed by MIPSS 104. In alternative embodiments, MIS 106 may provide a multimedia information stream to MIPSS 104. For example, MIS 106 may be a television receiver/antenna providing live television feed information to MIPSS 104. MIS 106 may be a video recorder providing the recorded video and/or audio stream to MIPSS 104. In alternative embodiments, MIS 106 may be a presentation or meeting recorder device that is capable of providing a stream of the captured presentation or meeting information to MIPSS 104. MIS 106 may also be a receiver (e.g., a satellite dish or a cable receiver) that is configured to capture or receive (e.g., via a wireless link) multimedia information from an external source and then provide the captured multimedia information to MIPSS 104 for further processing.

[63] Users may use user systems 102 to interact with the other systems depicted in Fig. 1. For example, a user may use user system 102 to select one or more multimedia documents and request MIPSS 104 to generate multimedia papers for the selected documents. Users may also use user systems 102 to view digital versions of the multimedia documents. For example, multimedia players executing on a user system may play multimedia information stored by a multimedia document. A user system 102 may be of different types including a personal computer, a portable computer, a workstation, a computer terminal, a network computer, a mainframe, a kiosk, a personal digital assistant (PDA), a communication device such as a cell phone, or any other data processing system.



[64] Output device 108 is capable of generating a multimedia paper document based upon the printable representation of the multimedia information received from MIPSS 104. Accordingly, output device 108 represents any device that is capable of outputting (e.g., printing, writing, drawing, imprinting, embossing, etc.) the printable representation of the multimedia information on a paper medium. For example, output device 108 may be a printer that is coupled to MIPSS 104. The printer may be configured to receive a signal from MIPSS 104 including a printable representation of multimedia information from MIPSS 104, and to generate a multimedia paper document based upon the printable representation of the multimedia information.

[65] According to an embodiment of the present invention, output device 108 may be incorporated as part of a multi-function device (or MFD) that is capable of performing a plurality of different functions in addition to allowing users to generate multimedia paper documents. For example, a MFD may allow users to copy, fax, or scan documents including multimedia paper documents. A MFD may also allow users to perform other functions. A MFD may also allow users to select multimedia documents for which printable representations are to be generated according to the teachings of the present invention. For example, a MFD may provide a user interface that allows a user to select one or more multimedia documents, request generation of printable representations for the selected multimedia documents, generate multimedia paper documents for the selected multimedia documents, and perform other functions such as copying, faxing, etc. on the printable representations or on the multimedia papers.

[66] Fig. 2A depicts a networked system including a MFD 200 according to an embodiment of the present invention. In the embodiment depicted in Fig. 2A, MFD 200 is coupled to MIPSS 104 that in turn is coupled to MIS 106. In the embodiment depicted in Fig. 2A, MIS 106 is a satellite dish or TV antenna that receives and provides multimedia information to MIPSS 104. MIPSS 104 generates a printable representation for the multimedia information. The printable representation may be forwarded to MFD 200 or may alternatively be stored by MIPSS 104.

[67] In the embodiment depicted in Fig. 2A, MFD 200 provides a user interface 202 that can be used by users to provide instructions or commands to MFD 200 and to view information output by MFD 200. Interface 202 comprises an area 204 that displays a list of documents 204-a including multimedia documents that can be selected by a user. The multimedia documents displayed in area 204 may be stored by MFD 200 or may be stored by other devices (such as MIPSS 104) coupled to MFD 200. In alternative embodiments, area

204 may display a list of documents accessible to MIPSS 104. The multimedia documents displayed in area 204 may correspond to television broadcast recordings, video clips, recorded meetings, etc.

[68] Area 204 may also display various details about the multimedia documents that are listed. In the embodiment depicted in Fig. 2A, for each multimedia document listed in area 204, the information displayed in area 204 includes information related to the date 204-c and time 204-b of the multimedia document recording. If a printable representation has already been generated for a multimedia document, the number of pages 204-d needed to print the printable representation (i.e., the number of pages in the multimedia paper document for the multimedia document) are also displayed. For example, the multimedia document titled "CNN/fn" stores a recording that was recorded on May 21, 2001 between 11:01AM and 1:00PM. A printable representation has been generated for the "CNN/fn" multimedia document and comprises 26 pages.

[69] A user may select one or more documents displayed in area 204 using an input device of MFD 104. In the embodiment depicted in Fig. 2A, the user may select a document by clicking on the document name in area 204 or alternatively by using "Select" button 206. For example, as shown in Fig. 2A, the user has selected a multimedia document titled "NewsHour" which corresponds to a news broadcast recorded on May 18, 2001 between 6:00-7:00PM. The user may then request MFD 200 to perform one or more functions provided by MFD 200 on the selected document(s). According to an embodiment of the present invention, the user may request generation of printable representations for the selected multimedia documents or may request generation of multimedia paper documents for the selected multimedia documents. The multimedia documents displayed in area 204 may also be indexed by MFD 200 that allows a user to perform familiar operations such as keyword searching, browsing for similar documents, etc. on the selected multimedia documents.

[70] User interface 202 provides a plurality of user interface features that allow a user to specify functions or operations to be performed on the selected document(s). For example, the user may select "Print" button 208 to instruct MFD 200 to print multimedia paper documents 210 for the multimedia documents selected by the user in area 204. According to an embodiment of the present invention, upon receiving a signal indicating selection of "Print" button 206 by a user, MFD 200 sends a signal to MIPSS 104 requesting generation of printable representations for the user-selected multimedia documents. If printable representations for the user-selected documents already exist, MIPSS 104

communicates the previously generated printable representations for the user-selected documents to MFD 200. Alternatively, if the printable representations do not exist, MIPSS 104 generates printable representations for the one or more user-selected documents and then provides the printable representations to MFD 200. MFD 200 may then generate (or print) multimedia paper documents for the user-selected documents based upon printable representations corresponding to the documents received from MIPSS 104. In alternative embodiments, MFD 200 may itself be configured to generate printable representations for user-selected multimedia documents.

[71] User interface 202 also provides a "Play" button 212 which when selected by a user causes MFD 200 to playback multimedia information from the user-selected multimedia document(s). For example, Fig. 2B depicts a user interface 214 that is displayed to the user upon selection of "Play" button 212 according to an embodiment of the present invention. Interface 214 allows the user to play back video and audio information contained in the "NewsHour" multimedia document selected by the user in area 204 of Fig. 2A. If MFD 200 is connected to one or more output devices (e.g., an output monitor, other networked output devices), the user may also select the output device to be used for the playback. For example, the user may indicate that the information is to be played back on the user's computer in the user's office (or on a television in a particular conference room, etc.). In specific embodiments of the present invention, the user may also indicate the time when the multimedia information is to be played back.

[72] Referring back to Fig. 2A, user interface 202 also provides a numeric keypad 216 that facilitates operations such as faxing of documents. For example, using keypad 216, a user may fax a multimedia paper document or a printable representation of a user-selected multimedia document to a recipient. The user may also make copies of the multimedia paper document by selecting "Copy" button 218. "Cancel" button 220 allows the user to cancel a pre-selected function.

[73] It should be apparent that MFD 200 and the user interfaces depicted in Figs. 2A and 2B are merely illustrative of an embodiment incorporating the present invention and do not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives. For example, in a networked environment, a web browser-enabled interface may be provided allowing a user to control the functions of MFD 200 from a remote location, for example, using the user's computer system or PDA, and the like.

[74] Fig. 3 is a simplified block diagram of a computer system 300 according to an embodiment of the present invention. Computer system 300 may be used as any of the computer systems depicted in Fig. 1. As shown in Fig. 3, computer system 300 includes at least one processor 302 that communicates with a number of peripheral devices via a bus subsystem 304. These peripheral devices may include a storage subsystem 306, comprising a memory subsystem 308 and a file storage subsystem 310, user interface input devices 312, user interface output devices 314, and a network interface subsystem 316. The input and output devices allow user interaction with computer system 300. A user may be a human user, a device, a process, another computer, and the like. Network interface subsystem 316 provides an interface to other computer systems and communication networks including communication network 110.

[75] Bus subsystem 304 provides a mechanism for letting the various components and subsystems of computer system 300 communicate with each other as intended. The various subsystems and components of computer system 300 need not be at the same physical location but may be distributed at various locations within network 100. Although bus subsystem 304 is shown schematically as a single bus, alternative embodiments of the bus subsystem may utilize multiple buses.

[76] User interface input devices 312 may include a keyboard, pointing devices, a mouse, trackball, touchpad, a graphics tablet, a scanner, a barcode scanner, a touchscreen incorporated into the display, audio input devices such as voice recognition systems, microphones, and other types of input devices. In general, use of the term "input device" is intended to include all possible types of devices and ways to input information using computer system 300.

[77] User interface output devices 314 may include a display subsystem, a printer, a fax machine, or non-visual displays such as audio output devices. The display subsystem may be a cathode ray tube (CRT), a flat-panel device such as a liquid crystal display (LCD), or a projection device. In general, use of the term "output device" is intended to include all possible types of devices and ways to output information from computer system 300.

[78] Storage subsystem 306 may be configured to store the basic programming and data constructs that provide the functionality of the computer system and of the present invention. For example, according to an embodiment of the present invention, software modules implementing the functionality of the present invention may be stored in storage subsystem 306 of MIPSS 104. For example, software modules that facilitate

generation of printable representations of the multimedia information may be stored in storage subsystem 306 of MIPSS 104. These software modules may be executed by processor(s) 302 of MIPSS 104. In a distributed environment, the software modules may be stored on a plurality of computer systems and executed by processors of the plurality of computer systems. Storage subsystem 306 may also provide a repository for storing various databases and files that may be used by the present invention. For example, the multimedia documents may be stored in storage subsystem 306. Storage subsystem 306 may comprise memory subsystem 308 and file storage subsystem 310.

[79] Memory subsystem 308 may include a number of memories including a main random access memory (RAM) 318 for storage of instructions and data during program execution and a read only memory (ROM) 320 in which fixed instructions are stored. File storage subsystem 310 provides persistent (non-volatile) storage for program and data files, and may include a hard disk drive, a floppy disk drive along with associated removable media, a Compact Digital Read Only Memory (CD-ROM) drive, an optical drive, removable media cartridges, and other like storage media. One or more of the drives may be located at remote locations on other connected computers.

[80] Computer system 300 itself can be of varying types including a personal computer, a portable computer, a workstation, a computer terminal, a network computer, a mainframe, a kiosk, a personal digital assistant (PDA), a communication device such as a cell phone, a game controller, or any other data processing system. Due to the ever-changing nature of computers and networks, the description of computer system 300 depicted in Fig. 3 is intended only as a specific example for purposes of illustrating the preferred embodiment of the computer system. Many other configurations of a computer system are possible having more or fewer components than the computer system depicted in Fig. 3. For example, several other subsystems may be included in computer system 300 depending upon the functions performed by system 300.

[81] Fig. 4 is a simplified high-level flowchart 400 depicting a method of generating a printable representation of multimedia information according to an embodiment of the present invention. The processing depicted in Fig. 4 may performed by MIPSS 104 (e.g., by software modules executing on MIPSS 104). In alternative embodiments of the present invention, the processing may be distributed among the various systems depicted in Fig. 1. Flowchart 400 depicted in Fig. 4 is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in

the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[82] As depicted in Fig. 4, according to an embodiment of the present invention, the method is initiated when MIPSS 104 receives a signal requesting generation of a printable representation for a multimedia document storing multimedia information (step 402). Alternatively, the signal received in step 402 may request generation of a multimedia paper document for a multimedia document. MIPSS 104 may receive the signal from a variety of different sources including a user system 102, a MFD 200, from an interface provided by MIPSS 104, from MIS 106, and the like. The signal may identify the multimedia document for which a printable representation is to be generated.

[83] In alternative embodiments of the present invention, the signal received in step 402 may comprise a stream of multimedia information (e.g., from MIS 106) for which a printable representation (or multimedia paper document) is to be generated. If the signal includes a multimedia information stream, MIPSS 104 may store the stream in a multimedia document and then generate a printable representation for the document. For purposes of explaining the processing in Fig. 4, it is assumed that the signal received in step 402 identifies a multimedia document for which a printable representation is to be generated.

[84] MIPSS 104 then accesses the multimedia document identified by the signal received in step 402 (step 404). The multimedia document identified by the signal received in step 402 may be stored by MIPSS 104 or may alternatively be stored by other devices or systems from where it can be accessed by MIPSS 104. In alternative embodiments of the present invention, the signal received in step 402 may itself comprise the multimedia document.

[85] MIPSS 104 then determines layout and format information to be used for generating the printable representation (step 406). The layout and format information specifies how the information is to be printed on the paper medium. For example, the layout and format information may comprise information identifying the paper-medium and size of the paper (e.g., letter size, legal size, A4 size, etc.) for which the printable representation is to be generated. The layout and format information may also identify special features of the paper (e.g., a paper with a letterhead, paper of a particular color, etc.) for which the printable representation is to be generated. In specific embodiments of the present invention, a default paper medium (e.g., letter size paper) may be selected for generating the printable representation of the multimedia document.

[86] Additionally, the layout and format information indicates the layout and formatting features to be used for generating the printable representation of the multimedia information. For example, according to an embodiment of the present invention, the layout and format information specifies how each type of information (e.g., audio, video, images, text, etc.) included in the multimedia information is to be printed. For example, for each type of information included in the multimedia information, the layout and format information may identify the area (or location or section of the paper medium) on the paper medium in which the information is to be printed, and the format or styles (e.g., font type, font size, bolding, underlining, number of columns per page, size of the columns, page margins, etc.) to be used for printing the information. In embodiments of the present invention which support multiple languages, the layout and format information may also indicate the language (or languages) to be used for printing the information. MIPSS 104 uses the layout and format information to generate the printable representation.

[87] For example for text information (e.g., CC text, text transcript of audio information) included in the multimedia information, the layout and format information may specify the font type and font size to be used for printing the text information, the number of columns to be used for printing the information, the size and location of the columns, the color of the font to be used for printing (which may depend on the color of the paper for which the printable representation is to be generated), line spacing, length of each line, number of words per line, bolding and capitalization techniques, and the like. The layout and format information may also identify the language to be used for printing the text information. For example, the layout and format information may indicate that the text is to be printed in two columns on each page with the English version in the first column and a Japanese translation of the English version in the second column.

[88] For audio information, the layout and format information may identify techniques to be used for converting the audio information to text information (i.e., techniques for generating a text transcript for the audio information), the format and styles for printing the audio transcript (which may be the same as for printing text information), and the like. For video information, the layout and format information may indicate how the video information is to be represented on paper. According to an embodiment of the present invention, the printable representation of the video information includes keyframes that are extracted from the video information. In this embodiment, the layout and format information may specify the sampling rate for extracting the keyframes, the number of keyframes that are



to be extracted from the video information, the order and placement of the keyframes on the paper medium, and other like information.

5 [89] Likewise, for other types of information included in the multimedia information, the layout and format information specifies the manner in which the multimedia information is to be printed on the paper medium. Accordingly, the layout and format information specifies how printable representations are to be generated for each type of information included in the multimedia information stored by the multimedia document.

[90] According to an embodiment of the present invention, the layout and format information is stored in the form of templates. The templates may be customized for a particular type of paper. For example, a first template may be defined for letter size paper and a second template different from the first template may be defined for A4 size paper. It should be apparent that one or more templates may be defined for each type and size of paper. If multiple templates are provided, the user may be allowed to select a particular template to be used for generating the printable representation. According to an embodiment of the present invention, information identifying a user-selected template may be included in the signal received in step 402. Default templates may also be specified. The user may also be allowed to create new templates, and to edit and modify previously configured templates. In this manner, the layout and format information is user-configurable.

25 [91] The goal of a template (or layout and format information in general) is to generate a printable representation which when printed on a paper medium generates a readable and comprehensible multimedia paper document. In order to create a readable version, the templates may adhere to many of the basic guidelines designed and used by the newspaper industry. For instance, the use of special fonts, multiple columns and shorter lines of text along with line spacing, bolding and capitalization techniques, and other type-setting features used by the newspaper industry may be specified in the templates. The layout and format information thus contributes towards making the multimedia paper document more readable and comprehensible.

30 [92] Figs. 5A and 5B depict a sample template according to an embodiment of the present invention. The template is defined using XML syntax but could easily be represented in other ways. The template is designed for use with letter-size (8.5 x 11 inch) sheets of white 24-lb. paper. As defined in the template, each sheet is configured to contain one title zone, two text zones, and a video zone. The title zone specifies the zone or area of the paper medium where the title is to be printed and the manner in which the title is to be printed. The first text zone specifies the zone or area or section of the paper medium where



the CC text included in the multimedia information is to be printed and the manner in which the CC text is to be printed. The second text zone specifies the zone or section of the paper medium where the Japanese translation of the CC text is to be printed and the manner in which the Japanese text is to be printed. It should be apparent that in alternative

5   embodiments of the present invention, CC text included in the multimedia information and which is a continuation of the information printed in the first text zone may be printed in the second text zone. The video zone specifies the zone or area of the paper medium where the video information included in the multimedia document is to be printed and the manner in which the video is to be printed.

[93]   The template information in Fig. 5A specifies that the title zone (identified by "ZONE\_ID 0") area is bounded by a rectangle whose left edge is located at a distance of 3 inches from the left margin of the page and whose right edge is located at a distance of 3 inches from the right margin of the page (i.e., the rectangle is 2.5 inches wide). The top edge of the title zone rectangle is located 0.75 inches from the top margin of the page and the bottom edge of the rectangle is located 9.6 inches from the bottom margin of the page (i.e., the rectangle is 0.65 inches high). The text in the title zone is configured to be English and is to be extracted from the header of the video clip. The title is configured to be printed in a 14 point, black Times font, and is centered within the title zone. The lines are to be single-spaced.

20   [94]   The first text zone (identified by "ZONE\_ID 1") is also bounded by a rectangle whose left edge is located at a distance of 1.1 inches from the left margin of the page, whose right edge is located at a distance of 5.4 inches from the right margin of the page, whose top edge is located 1.5 inches from the top margin of the page, and whose bottom edge is located 1.0 inch from the bottom margin of the page. The text in the first text  
25   zone is to be printed in the English language. The origin of the text to be printed is defined to be CC text included in the multimedia information. The text is to be printed in a black 10 point Garamond font. Lines are to be single-spaced. Subject changes in the closed caption (which are usually indicated in the CC text by three greater-than signs ">>>") are to be shown by inserting a 1.5 line break and by bolding the first three words. Speaker changes  
30   (which are usually indicated in CC text by two greater-than signs ">>") are to be shown with a single-line break with no emphasis. Annotations (that indicate words in the transcript that occur in a user's profile) (described below in more detail) are to be shown with italicized text and blue underlining.

[95] The second text zone (identified by "ZONE\_ID 2") is also bounded by a rectangle whose left edge is located at a distance of 4.6 inches from the left margin of the page, whose right edge is located at a distance of 1.9 inches from the right margin of the page, whose top edge is located 1.5 inches from the top margin of the page, and whose  
5 bottom edge is located 1.0 inch from the bottom margin of the page. Unlike the first zone, the text in the second text zone is to be printed in Japanese. A translation source to facilitate the translation to Japanese is identified. The text is to be printed in a black 10 point AsianGaramond font. Lines are to be single-spaced. Subject changes in the closed caption text (which are usually indicated in CC text by three greater-than signs ">>>") are to be shown by inserting a 1.5 line break and by bolding the first three words. Speaker changes (which are usually indicated in CC text by two greater-than signs ">>") are to be shown with a single-line break with no emphasis. Annotations to words or phrases are to be shown with italicized text and blue underlining. Further details related to annotations are provided below.

[96] The video zone (identified by "ZONE\_ID 3") is also bounded by a rectangle whose left edge is located at a distance of 3.2 inches from the left margin of the page, whose right edge is located at a distance of 4.5 inches from the right margin of the page, whose top edge is located 1.5 inches from the top margin of the page, and whose bottom edge is located 1.0 inch from the bottom margin of the page. The source for the displayed data in the video zone is to be a set of keyframes that are generated by sampling the video channel of the multimedia information at a rate of 1 frame per second. Text in those frames is to be annotated by drawing a red box around it with line-width of 3-points. The keyframes are to be divided into sets of four. Each set is to be 0.4 inches wide and 0.3 inches high. The keyframes from each set are to be laid out within the video zone by sequentially packing them into the available space. Each group of four keyframes is to be annotated with  
25 an interleaved 2-of-5 barcode 0.8 inches wide and 0.15 inches high that appears underneath the group.

[97] It should be apparent that the template depicted in Figs. 5A and 5B is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would  
30 recognize other variations, modifications, and alternatives.

[98] Referring back to Fig. 4, MIPSS 104 then generates a printable representation of the multimedia information stored in the multimedia document accessed in step 402 based upon the layout and format information determined in step 406 (step 408). Generating a printable representation for the multimedia document involves generating a

printable representation for each type of information included in the multimedia information based upon the layout and format information.

[99] If the signal received in step 402 requested generation of a multimedia paper document, MIPSS 104 may then print the printable representation of the multimedia information to generate the multimedia paper document (step 410). Alternatively, MIPSS 104 may communicate the printable representation of the multimedia information generated in step 408 to an output device 108 (e.g., a printer, a MFD, etc.) that is configured to generate the multimedia paper document (step 412). Other operations may also be performed on the printable representation of the multimedia information (step 414). For example, the printable representation may be stored for future generation of multimedia paper document, the information may be faxed, the information may be searched, indexed, annotated, etc., and the like.

[100] Fig. 6 is a simplified high-level flowchart depicting processing performed in step 408 of Fig. 4 according to an embodiment of the present invention. The processing depicted in Fig. 6 may be performed by software modules executing on MIPSS 104, by hardware modules coupled to MIPSS 104, or a combination thereof. In alternative embodiments of the present invention, the processing may be distributed among the various systems depicted in Fig. 1. The processing depicted in Fig. 6 is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[101] As described above, in step 408 MIPSS 104 generates a printable representation of the multimedia information based upon the layout and format information determined in step 406. As part of the processing performed in step 408, MIPSS 104 divides or indexes the multimedia information contained by the multimedia document into sequential segments or portions of a particular time length (step 602). Each segment is characterized by a starting time and an ending time. Each segment comprises multimedia information occurring between the starting time and ending time associated with the segment. In other words, each segment or section comprises multimedia information for a specific time period. A sequential list of segments represents the entire multimedia information stored in the multimedia document. For example, according to an embodiment of the present invention, a 10-second time period may be used for segmenting the multimedia information. Using the 10-second time period value, a 5-minute video recording may be divided into 30 segments or sections. The first segment comprises multimedia information for the first 10 seconds of the

multimedia document, the second segment comprises multimedia information for the next 10 seconds, the third segment comprises multimedia information for the next 10 seconds, and so on. The value of the time period to be used for segmenting the multimedia document may be user-configurable.

5           [102] From the segments generated in step 602, MIPSS 104 then selects a set of segments or portions of the multimedia document comprising multimedia information that is to be included in the printable representation of the multimedia information (step 604). According to an embodiment of the present invention, all the segments generated in step 602 are selected to be included in the printable representation. According to other embodiments of the present invention, a subset of the segments generated in step 602 may be selected for inclusion in the printable representation of the multimedia information based upon some selection criteria. The selection criteria may be user configurable.

10           [103] According to one such embodiment, MIPSS 104 may compare multimedia information stored by successive segments and only select those segments for inclusion in the printable representation that contain additional information relative to their preceding segment. In this manner, segments comprising repetitive or redundant information are not selected for inclusion in the printable representation. For example, there may be periods of time within a video recording wherein the audio or video content information does not change (e.g., during a period of "silence" or "blankness" on the video recording). Segments comprising multimedia information corresponding to such periods of time may not be selected by MIPSS 104 in step 604.

          [104] According to another embodiment of the present invention, MIPSS 104 may select only those segments for inclusion in the printable representation that contain information relevant to the user who has requested generation of the printable representation.

25           For example, MIPSS 104 may select only those segments for inclusion in the printable representation that contain multimedia information related to user-specified topics of interest (which may be specified in a user profile). For example, a user may have specified an interest in all information related to topic "Afghanistan." In this embodiment, MIPSS 104 may scan the multimedia information contained by the various segments and select only  
30           those segments for inclusion in the printable representation that contain information related to Afghanistan. Various techniques known to those skilled in the art may be used to facilitate selection of the segments based upon their content and their relevance to user-specified topics.

[105] According to another embodiment of the present invention, MIPSS 104 may apply a summarization technique to select segments to be included in the printable representation. Applying the summarization technique, only those segments that satisfy some selection criteria may be selected for inclusion in the printable representation. For example, 5 for a multimedia document corresponding to an audio recording, MIPSS 104 may only select those segments for inclusion that contain the first sentence spoken by each speaker (or alternatively segments that contain the first line of each paragraph of CC text). This reduces the size of the printable representation and as a result reduces the number of pages needed to print the printable representation. Various other techniques known to those of skill in the art may also be used to determine which segments are to be included in the printable representation of the multimedia information.

[106] MIPSS 104 then paginates the segments (i.e., determines on which page a particular segment is to be printed) selected in step 604 (step 606). According to an embodiment of the present invention, for each page starting with the first page, MIPSS 104 determines the segments to be printed on the page based upon the layout and format information which influences the amount of information that can be printed on a page. In this manner, MIPSS 104 determines the amount of multimedia information to be printed on each page and the total number of pages required to print the multimedia information stored in the multimedia document. For each page, MIPSS 104 determines the start time for information printed on the page (corresponding to the start time of the first segment printed on the page) and the end time for information printed on the page (corresponding to the end time of the last segment printed on the page).

[107] The number of segments that can be printed on a particular page is influenced by the layout and format information and the contents of the segments. The size of the contents of each segment is in turn influenced by the time period used for segmenting the multimedia information stored in the multimedia document. For example, the amount of information stored by a segment is generally directly proportional to the value of the time period used for segmenting the multimedia document. 25

[108] According to an embodiment of the present invention, for a given template storing the layout and format information and for a particular segmentation time period, the number of segments printed on a particular page is fixed for each page of the multimedia paper document. For example, based upon a particular template and a particular time period used for segmenting the multimedia information, MIPSS 104 may determine that multimedia information corresponding to "M" segments (where  $M > 0$ ) can be printed on 30

each page of the multimedia paper document. Based upon this segments-per-page value, MIPSS 104 may then determine the total number of pages in the multimedia paper document and the segments to be printed on each page.

[109] For example, for a 5-minute video recording which is divided into 30 segments using a 10-second segmentation value, and assuming that all segments are selected for inclusion in the printable representation in step 604, MIPSS 104 may determine that multimedia information corresponding to 12 segments will be printed per page of the multimedia paper document. Using this segments-per-page value, MIPSS 104 may determine that 3 pages ( $\left\lceil \frac{30}{12} \right\rceil = 3$ ) will be needed to print the multimedia information (i.e.,

the multimedia paper document will contain 3 pages). Multimedia information corresponding to segments 1-12 will be printed on the first page of the multimedia paper document, multimedia information corresponding to segments 13-24 will be printed on the second page of the multimedia paper document, and multimedia information corresponding to segments 25-30 will be printed on the last or third page of the multimedia paper document.

[110] In alternative embodiments of the present invention, the number of segments printed on a particular page may vary from page to page of the multimedia paper document based upon the contents of the segments. In this embodiment, the number of segments to be printed on a particular page is influenced by the type and contents of multimedia information contained in the segments selected in step 604. In this embodiment, for each page, starting with the first page of the multimedia paper document, MIPSS 104 determines the number of selected segments (starting with the segment having the earliest starting time) which can be printed on each page. In this manner the number of segments that can be printed on a page are determined on a sequential page-per-page basis starting with the first page.

[111] For example, for a 5-minute video recording which is divided into 30 segments using a 10-second segmentation value and assuming that all the segments are selected for inclusion in step 604, MIPSS 104 may determine that multimedia information corresponding to segments 1-10 can be printed on the first page, multimedia information corresponding to segments 11-25 can be printed on the second page of the multimedia paper document, and multimedia information corresponding to sections 25-30 can be printed on the third page of the multimedia paper document. Accordingly, in this embodiment of the present invention, the number of segments printed on each page of the multimedia document

may vary from page to page based upon the contents of the segments to be printed. Various other techniques may also be used to determine how the selected segments will be printed.

5 [112] MIPSS 104 then generates a printable representation for each page determined in step 606 based upon the layout and format information (step 608). As part of step 608, for each page, MIPSS 104 determines segments associated with that page, and generates printable representations for the various types of information included in the multimedia information corresponding to the segments. Various different techniques may be used by MIPSS 104 to generate printable representation for the various types of information included in the multimedia information.

[113] For example, for CC text information included in the multimedia information, MIPSS 104 may apply the formatting styles specified by the layout and format information. For audio information, MIPSS 104 may generate a text transcript for the audio information by applying audio-to-text conversion techniques (which may also be specified in the layout and format information) and then apply the text formatting. For video information, MIPSS 104 may apply various keyframe extraction techniques (which may be specified in the layout and format information) to extract keyframes from the video information included in the selected segments of the multimedia information. According to an embodiment of the present invention, MIPSS 104 extracts keyframes that capture salient features of the video information (or keyframes that are informative) for a particular segment of the multimedia information. For example, images of faces are often quite informative. In choosing keyframes for a news broadcast, MIPSS 104 may select keyframes whose contents are different from the anchorperson. This increases the information conveyed by the keyframes.

25 [114] Several other techniques known to those of skill in the art may also be applied by MIPSS 104 to generate a printable representation for the multimedia information. For example, the article "Key frame selection to represent a video," written by Frederic Dufaux and published in the Proceedings of the International Conference on Image Processing, Vancouver, 2000, describes techniques for selecting keyframes for representing a video. The entire contents of this article are herein incorporated by reference in their entirety for all purposes.

30 [115] The printable representation of the multimedia information generated by MIPSS in step 408 may then be printed to generate a multimedia paper document, communicated to a device capable of generating a multimedia paper document, or subjected to other operations according to step 410, 412, or 414 depicted in Fig. 4.



[116] Fig. 7A depicts a page 700 from a multimedia paper generated according to an embodiment of the present invention for a multimedia document. In the embodiment depicted in Fig. 7A, the multimedia document corresponds to a television broadcast recording. As depicted in Fig. 7A, page 700 comprises a title section 702, a first text section 704, a second text section 706, a video section 708, and a controls section 710. Page 700 depicted in Fig. 7A is merely illustrative of a multimedia paper document page according to an embodiment of the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[117] Page 700 depicted in Fig. 7A is imprinted with multimedia information corresponding to ten segments. According to an embodiment of the present invention, identifiers 712 identifying the segments are displayed in text sections 702 and 704, and in video section 708. The segment identifiers are printed proximally close to information corresponding to the respective segments. Page 700 also displays time span information 714 that indicates the start time and end time corresponding to information printed on page 700. For example, the information printed on page 700 represents multimedia information recorded during the first 5:29 minutes of the recording. The page number 715 for each page is also displayed. Accordingly, page 700 depicted in Fig. 7A is the first page of the multimedia paper document.

[118] As shown in Fig. 7A, title section 702 displays title information for the multimedia paper document. As depicted in Fig. 7A, the title information includes information identifying the source 716 of the multimedia information recording. According to an embodiment of the present invention, source information 716 corresponds to the name (e.g., filename) of the multimedia document for which the multimedia paper document has been generated. The title information displayed in section 702 also includes the time 718 when the multimedia document was recorded, the total time length 720 of the recording, and the date 722 of the recording. For example, page 700 is the first page from a multimedia paper document generated for "CNN News site (Channel 203)" television program which was recorded on May 30, 2001 starting at 12:59PM and has a total length of 56 minutes and 40 seconds.

[119] Text sections 704 and 706 display text information included in the multimedia document for which the multimedia paper document has been generated. In the embodiment depicted in Fig. 7A, text sections 704 and 706 display CC text included in the multimedia information. In alternative embodiments of the present invention, text sections



704 and 706 may display a transcript of the audio information included in the multimedia information.

[120] Identifiers 712 are printed in (or next to) to the text sections.

5 According to an embodiment of the present invention, each identifier 712 printed on a page of the multimedia paper document identifies a segment of the multimedia document that is printed on the page. The segment identifiers are printed proximally close to information corresponding to the respective segments.

[121] According to alternative embodiments of the present invention, identifiers 712 represent time points in the multimedia document. In this embodiment, identifiers 712 are printed proximal to information printed on a page that occurs close to the time corresponding to the identifier. For example, an identifier 712 printed on a particular page may represent a time in the time span for the particular page. For example, if the time span for a particular page is 0:00min-5:29min (e.g., time span of page 700 depicted in Fig. 7A), a particular identifier 712 may represent a time of 3:00min, i.e., 3 minutes into the multimedia recording. The particular identifier is printed proximal to information that occurs at a time of 3 minutes into the multimedia recording.

[122] In the embodiment depicted in Fig. 7A, text sections 704 and 706 display the CC text in the English language. However, in alternative embodiments of the present invention that support multiple languages, the text may be printed in various languages or combinations thereof. The languages used to print the text may be different from the language of the CC text included in the multimedia information or the language of the audio information included in the multimedia information. For example, the CC text associated with a video broadcast recording may be in English, but the text corresponding to the CC text printed on the multimedia paper document may be in a different language, for  
25 example, in Japanese (see Fig. 7D). Various different formats and styles may be used to print text in the various languages. For example, according to an embodiment of the present invention, English text may be printed in text section 704 depicted in Fig. 7A and the corresponding translated Japanese text may be printed in text section 706. In alternative  
30 embodiments, each line of English text printed in a text section may be followed by a Japanese translation of the text, and the like. Various other formats may also be used to print text in different languages. The translation of text from one language to another may be performed by MIPSS 104 or alternatively may be performed by some other service or application and then provided to MIPSS 104.

[123] The present invention also takes advantage of the automatic story segmentation that is often provided in close-captioned (CC) text from broadcast news. Most news agencies who provide CC text as part of their broadcast use a special syntax in the CC text (e.g., a ">>>" delimiter to indicate changes in story line or subject, a ">>" delimiter to indicate changes in speakers, etc.) to indicate the beginning of a new story. Given the presence of this kind of information in the CC text transcript, the present invention can further enhance the contents of the paper document with newspaper layout techniques such as bolding and line spacing that typically signify a new story. For example, as depicted in Fig. 7A, the first line of each new story is bolded. Further, additional spacing is provided between text portions related to different story lines to clearly demarcate the different stories. This further enhances the readability of the multimedia paper document.

[124] For each speaker identified in the CC text, information related to the speaker may be printed on page 700 (not shown). The information may include a name of the speaker, an address associated with the speaker, the title (e.g., CEO, etc.) of the speaker, and other like information related to or identifying the speaker. The information may also include information printed on a business card of the speaker. The information related to the speakers may be determined from multimedia information stored in the multimedia document or may alternatively be determined from other information resources available to MIPSS 104.

[125] According to an embodiment of the present invention, video section 708 displays keyframes extracted from the video information corresponding to the CNN News Site (Channel 203) news recording. As depicted in Fig. 7A, four keyframes have been extracted from the video information for each segment and displayed in video section 706. Identifiers 712 are printed in the upper right hand corner of each set of four keyframes. As described above, according to an embodiment of the present invention, identifiers 712 identify the segments from which the keyframes have been extracted. In alternative embodiments of the present invention, identifiers 712 may represent specific time points within the multimedia recording. The number of keyframes that are extracted from each segment and the number of keyframes printed on each page of the multimedia paper document for each segment are user configurable. For example, according to one embodiment of the present invention, only one keyframe may be displayed for each segment, and the like. As previously stated, several different keyframe extraction techniques known to those of skill in the art may be used to extract keyframes from the video information included in the multimedia information. Additionally, several different techniques known to those of skill in the art may be used to display one or more of the extracted keyframes.

[126] As shown in Fig. 7A, identifiers 712 are printed in the text sections and also in the video section. A user may thus use identifiers 712 to correlate a portion of text printed in text sections 704 or 706 with a set of keyframes displayed in video section 708, and vice versa. For example, while a user is skimming the text section, the user may read a particular portion of text proximal to a particular identifier and locate keyframes related to or co-occurring with the particular portion of text using the particular identifier. Alternatively, the user may see an identifier for a particular keyframe (or set of keyframes) and use the identifier to locate text that describes what is being talked about at about the time that the keyframe(s) appeared in the video information. Identifiers 712 thus provide a sort of visual reference as well as a context for reading the text and the keyframes. This enhances the readability of the multimedia paper document.

[127] User-selectable identifiers 726 are printed on page 700. In the embodiment depicted in Fig. 7A, user-selectable identifiers 726 are printed as barcodes. A barcode 726 is printed for each segment printed on page 700. For example, barcode 726-1 corresponds to segment 1, barcode 726-2 corresponds to the segment 2, barcode 726-3 corresponds to the segment 3, and so on. In alternative embodiments of the present invention, various other techniques, besides barcodes, may be used to represent the user-selectable identifiers. As will be discussed below in further details, user-selectable identifiers 726 provide a mechanism for the reader of the multimedia paper document to access or retrieve multimedia information using the multimedia paper document.

[128] In alternative embodiments of the present invention where identifiers 712 represent specific time points in the multimedia information recording, barcodes 726 may be correlated to identifiers 712 and may also correspond to specific time points in the multimedia information recording. According to an embodiment of the present invention, barcodes 726 may correspond to the same time points as identifiers 712. Further details related to user-selectable identifiers 726 are provided below. User-selectable identifiers 726 are printed in a manner that does not reduce or affect the overall readability of the multimedia paper document.

[129] As depicted in Fig. 7A, controls section 710 displays a plurality of barcodes 724 corresponding to controls that may be used by a user to control playback of multimedia information corresponding to user-selected segments. Further details related to controls section 710 are provided below. Barcodes 724 are printed in a manner that does not reduce or affect the overall readability of the multimedia paper document.

[130] Fig. 7B depicts a second page 750 that follows page 700 depicted in Fig. 7A in a multimedia paper document according to an embodiment of the present invention. Title section 702 is not displayed on page 750. Page 750 displays text and keyframes corresponding to 11 segments (as compared to page 700 wherein information corresponding to 10 segments is displayed) of the multimedia document. The information displayed on page 750 corresponds to multimedia information corresponding to 5:29 minutes through 11:30 minutes of the recording (as indicated by time span information 714).

[131] Fig. 7C depicts a page 760 from a multimedia paper document generated according to an embodiment of the present invention for a multimedia document. Page 760 depicted in Fig. 7C corresponds to a page from a multimedia paper document generated for multimedia information recorded during a meeting. Information identifying the meeting is printed in title section 766. As depicted in Fig. 7C, page 760 comprises a first text section 762, a second text section 764, a video section 768, and a controls section 770.

[132] Closed-caption text (or a text transcript of the audio information) included in the multimedia document is printed in text section 762. A Japanese translation of the text printed in text section 762 is printed in text section 764. This is different from pages 700 and 750 depicted in Figs. 7A and 7B, respectively, wherein CC text was printed in both the text sections. For example, in Fig. 7A, the CC text printed text section 706 is a continuation of the text printed in text section 704. Various translation resources may be used to generate the Japanese translation printed in section 764 of Fig. 7C. It should be apparent that in alternative embodiments, the CC text may be translated to other languages and printed in a multimedia paper document.

[133] Page 760 depicted in Fig. 7C is merely illustrative of a multimedia paper document page according to an embodiment of the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[134] Given a multi-paged multimedia paper document comprising pages of the type depicted in Figs. 7A, 7B, or 7C, a reader can quickly skim the contents of the multimedia paper document to see if anything relevant might be present in the multimedia information for which the multimedia paper document was generated. The time required to skim and comprehend information printed in the multimedia paper document will be much smaller than the time the user would otherwise have to spend viewing the multimedia information (e.g., new broadcast recording). The present invention thus allows the user to save valuable time when "reading" multimedia information.

[135] Fig. 8A depicts a page 800 from a multimedia paper document generated for a recorded meeting according to an embodiment of the present invention. Page 800 depicted in Fig. 8A is merely illustrative of a page from a multimedia paper document and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[136] The recorded meeting for which page 800 is generated may store multimedia information that includes video information, audio information, slides information, and whiteboard information. Techniques for recording meetings has been described in U.S. Non-Provisional Patent Application No. 09/728,560, filed November 30, 2000, and U.S. Non-Provisional Patent Application No. 09/728,453, filed November 30, 2000.

[137] The slides information included in a recorded meeting may comprise information related to slides (e.g., a PowerPoint presentation slides) presented by the presenter during the meeting. The whiteboard information may comprise information related to text and drawings drawn on a whiteboard during the meeting. Accordingly, in addition to text (which may correspond to a transcript of the audio information) and video information, slides information and whiteboard information are also included in the printable representation of the recorded meeting. The text, video, slides, and whiteboard information may then be printed on a paper medium as depicted in Fig. 8. Accordingly, the text information is printed in sections 802 and 804, video information is printed in section 806, and slides 806 and whiteboard images 808 are printed inline with the text sections.

[138] According to an embodiment of the present invention, during generation of the printable representation for the recorded meeting, MIPSS 104 synchronizes the slides information and whiteboard information with the audio and video information using timestamps associated with the various types of information. When the multimedia information corresponding to the recorded meeting is divided into segments, each segment may comprise text information, video information, slides information, and whiteboard information. When the multimedia paper document is generated, one or more slides are then printed in close proximity to the identifier of a segment that contains slides information related to the printed slides. The slides are thus printed close to when they were presented. Likewise, images of the whiteboard are printed in close proximity to the identifier of a segment that contains the whiteboard information. The whiteboard images are thus printed close to when they were presented. In the embodiment depicted in Fig. 8A, the slides and whiteboard images are printed inline with the text sections.

[139] Various different layout and format guidelines may be used for printing the slides and whiteboard information. For example, in Fig. 8B, slides 808 and whiteboard images 810 are printed in the margins of the multimedia paper document next to text sections 802. Fig. 8C shows yet another layout pattern for printing the slides and whiteboard information. In Fig. 8C, the slides and whiteboard images are superimposed on video keyframes belonging to the segments to which the slides and whiteboard images belong.

[140] As described above, audio information included in multimedia information stored by a multimedia document is displayed in the form of a text transcript of the audio information. According to an embodiment of the present invention, various other features of the audio signal included in the multimedia information may also be represented in the printable representation of the multimedia document. According to an embodiment of the present invention, visual markers are used to represent the various features and when printed on a paper medium improve the overall readability and understandability of the multimedia paper document.

[141] For example, Fig. 9A depicts a page 900 from a multimedia paper document displaying visual markers to denote various attributes of the audio information or of the CC text information included in the multimedia information for the multimedia document for which the multimedia paper document is generated according to an embodiment of the present invention. Page 900 depicted in Fig. 9A is merely illustrative of a multimedia paper document page according to an embodiment of the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[142] As depicted in Fig. 9A, a gap or white space 908 is shown in text section 904 corresponding to one more segments that do not contain any CC text information but may comprise other types of information (e.g., video information) which is printed on the page (e.g., keyframes in video section 906). The gap may represent a section of the recording wherein there is no audio or CC text. Alternatively, the gap may represent a section of the recording where there is no CC text and the audio information during the section cannot be translated to text. For example, someone is speaking in a foreign language for which an English translation is not available. The length of gap 908 may be proportional to the length of the empty CC text or absence of audio information. If the multimedia information does not include both the audio information and the CC text for a specified time period, a visual marker such as "SILENCE" may be printed in gap 908.

10033894 1434

[143] The video information during a gap in the CC text or audio information may still contain important information and is thus displayed as keyframes in video section 906. For example, if someone is speaking in a foreign language for which an English translation is not available for display in the text section, the video during this period may display text (e.g., subtitles) corresponding to what is being spoken. Accordingly, keyframes displaying the text may be printed in video section 906 while a gap is printed in the text section. According to one embodiment of the present invention, the text images may be extracted from the video keyframes during the gap period and printed in gap 908. For example, as depicted in Fig. 9B, text images 920 have been extracted from video keyframes corresponding to the gap period, and the extracted text images 920 are printed in gap space 908. According to yet another embodiment of the present invention, optical character recognition (OCR) techniques may be applied to the video keyframes for the gap period and the results of the OCR may be printed in gap space 908. For example, as depicted in Fig. 9C, OCR techniques have been applied to the video keyframes during the gap period, and the resultant OCR'd text 930 (which may contain spelling errors) is printed in gap 908.

[144] Other features of the audio information may also be represented via visual markers printed on the multimedia paper document. For example, features of audio information such as people singing, multiple people talking at the same time, people arguing, speaking in soothing tones, significant increases in audio volumes, periods of silence (described above) etc. can be identified and represented in the multimedia paper document using visual markers. For example, as depicted in Fig. 9A, visual markers "(Singing)" 910 are printed where the audio information contains people singing. The visual markers thus make it easy for the reader of the multimedia paper document to quickly locate such parts of the audio in the multimedia document.

[145] Several different techniques known to those skilled in the art may be used to identify special features of the audio information in the multimedia information. The following references discuss a few techniques that may be used to identify features of audio signals. The entire contents of the following references are herein incorporated by reference for all purposes:

(1) L.S. Chen, H. Tao, T.S. Huang, T. Miyasato, R. Nakatsu, "Emotion Recognition from Audiovisual Information," Proc. IEEE Workshop on Multimedia Signal Processing, Los-Angeles, CA, USA, pp. 83-88, 1998;



(2) K. Sonmez, L. Heck, M. Weintraub, "Multiple Speaker Tracking and Detection: Handset Normalization and Duration Scoring," Digital Signal Processing, 10(1/2/3), 133-143, 2000; and

(3) F. Dellaert, T. Polzin, A. Waibel, "Recognizing emotion in speech." Proceedings ICSLP 96. Fourth International Conference on Spoken Language Processing (Cat. No.96TH8206). IEEE. Vol.3, pp.1970-1973, 1996. New York, NY, USA.

[146] As described above, video information included in multimedia information stored by a multimedia document is displayed in the form of one or more keyframes extracted from the video information and printed on the multimedia paper document. According to an embodiment of the present invention, various other features of the video information included in the multimedia information may also be represented in the printable representation of the multimedia document. According to an embodiment of the present invention, visual markers may be used to represent the various features of the video information and when printed on a paper medium improve the overall readability and understandability of the multimedia paper document.

[147] For example, features that can be recognized from video information may include faces, facial expressions of speakers depicted in the video (e.g., a facial expression indicating anger), recognition of speakers, hand gestures, logos or signs displayed in the video, presence of certain buildings or geographical locations, meetings, animals, crowds, and the like. According to an embodiment of the present invention, these features are recognized from the video information and represented in the multimedia paper documents using visual markers. For example, expressions (e.g. "Anger," "Laughter", etc.), geographical locations, special building names, etc. can be shown with a text-based annotation next to the corresponding video keyframe. Speaker face recognition results may be shown with the name of the speaker printed next to a keyframe depicting the speaker. Logos and signs recognized from the video can be displayed or the names of companies corresponding to the logos may be displayed.

[148] Several different styles and formats may be used to display the visual markers for the video features. According to an embodiment of the present invention, the layout and format information may specify the styles or formats are to be used. For example, the visual markers for the video information may be displayed close to keyframes corresponding to the visual markers, or in the margins of the multimedia paper document. In alternative embodiments, the visual markers may be displayed interspersed with the text section information. Different types of fonts may be used to display the visual markers. The



visual markers thus make it easy for the reader of the multimedia paper document to quickly locate such parts of the video in the multimedia document.

[149] The text displayed in the text sections of a multimedia paper document (e.g., text sections 704 and 706 depicted in Fig. 7A) may also be modified based upon  
5 recognition of certain video features. For example, text printed in the text section and spoken by a particular speaker may be highlighted, and the like.

[150] Several different techniques known to those skilled in the art may be used to identify special features of the video information in the multimedia information. The following references discuss a few techniques that may be used to identify features of the video data. The entire contents of the following references are herein incorporated by reference for all purposes:

(1) A. Essa, A. P. Pentland, Coding Analysis Interpretation and Recognition of Facial Expressions, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 757-763, 1997;

(2) G. Donato, M. Bartlett, J. Hager, P. Ekman, and T. Sejnowski, Classifying Facial Actions, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 21, no. 10, pp. 974-989, Oct. 1999;

(3) A.F. Bobick, A.D. Wilson, A State based approach to the representation and recognition of gesture, IEEE Trans. on Pattern Analysis and Machine Intelligence, pp. 1325-1337, 1997;

(4) H. A. Rowley, S. Baluja, T. Kanade, "Neural network-based face detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 20, no. 1, 23-38, 1998;

(5) D. S. Doermann, E Rivlin, and I. Weiss. Applying algebraic and differential  
25 invariants for logo recognition. Machine Vision and Applications, 9(2):73-86, 1996;

(6) H. Li, D. Doermann, and O. Kia. Automatic Text Detection and Tracking in Digital Video. IEEE Transactions on Image Processing - Special Issue on Image and Video Processing for Digital Libraries, 9(1), pages 147-156, 2000;

(7) P. Suda, C. Bridoux, B. Kammerer, G. Manderlechner, "Logo and word  
30 matching using a general approach to signal registration," Fourth International Conference on Document Analysis and Recognition, Ulm, Germany, August 18-20, 1997, 61-65;

(8) H. Li, D. Doermann, and O. Kia. Text Extraction and Recognition in Digital Video. Proceedings of Third IAPR Workshop on Document Analysis Systems, pages 119-128, 1998;



highlighted. In this manner, the multimedia paper document provides a convenient tool that allows a user to readily locate portions of the multimedia paper document that are relevant to the user. Since the multimedia paper document comprises a printable representation of multimedia information, the multimedia paper document generated according to the teachings  
5 of the present invention allows the user to identify portions of multimedia information that are of interest to the user.

[155] According to an embodiment of the present invention, information specifying topics of interest to the user may be stored in a user profile. One or more words or phrases may be associated with each topic of interest. Presence of words and phrases  
10 associated with a particular user-specified topic of interest indicates presence of information related the particular topic. For example, a user may specify two topics of interest--"George Bush" and "Energy Crisis". Words or phrases associated with the topic "George Bush" may include "President Bush," "the President," "Mr. Bush," and other like words and phrases. Words or phrases associated with the topic "Energy Crisis" may include "industrial pollution," "natural pollution," "clean up the sources," "amount of pollution," "air pollution", "electricity," "power-generating plant," and the like. Probability values may be associated with each of the words or phrases indicating the likelihood of the topic of interest given the presence of the word or phrase. Various tools may be provided to allow the user to configure topics of interest, to specify keywords and phrases associated with the topics, and to specify probability values associated with the keywords or phrases.

[156] According to an embodiment of the present invention, after generating a printable representation of multimedia information stored in a multimedia document (in step 408 of Fig. 4), MIPSS 104 accesses the user profile information and determines topics of interest specified in the user profile and keywords and phrases associated with the topics of  
25 interest. MIPSS 104 then searches the printable representation of the multimedia information to identify locations within the printable representation of words or phrases associated with the topics of interest. As described above, presence of words and phrases associated with a particular user-specified topic of interest indicates presence of the particular topic relevant to the user. According to one embodiment of the present invention, MIPSS 104 searches the  
30 text sections included in the printable representation of a multimedia document to locate words or phrases associated with the user topics. If MIPSS 104 finds a word or phrase in the printable representation that is associated with a topic of interest, the word or phrase is annotated in the printable representation. Several different techniques may be used to annotate the word or phrase. For example, the word or phrase may highlighted, bolded,

underlined, demarcated using sidebars or balloons, font may be changed, etc. The annotations are then printed on the multimedia paper document generated by printing the annotated printable representation of the multimedia information.

5 [157] According to an embodiment of the present invention, MIPSS 104 may also highlight keyframes (representing video information of the multimedia document) related to user specified topics of interest. According to an embodiment of the present invention, MIPSS 104 may use OCR techniques to extract text from the keyframes included in the printable representation of the multimedia information. The text output by the OCR techniques may then be compared with words or phrases specified in a user's profile. If there is a match, the keyframe corresponding to the matched word or phrase (i.e., the keyframe from which the matching word or phrase was extracted) may be annotated in the printable representation. Several different techniques may be used to annotate the keyframe. For example, a special box may surround the keyframe, the matching text in the keyframe may be highlighted or underlined or displayed in reverse video, and the like. The keyframe annotations are then printed on the multimedia paper document generated by printing the annotated printable representation of the multimedia information.

10001891-11991

25 [158] According to another embodiment of the present invention, MIPSS 104 may identify information stored by the multimedia document that is relevant to user-specified topics of interest even before the printable representation for the multimedia document has been generated. In this embodiment, MIPSS 104 analyzes the multimedia information stored in the multimedia document to identify information relevant to user-specified topics of interest. For example, MIPSS 104 may analyze the video information contained in the multimedia document to identify video frames that contain information relevant to user-specified topics of interest. Various different techniques, e.g., OCR techniques, known to those skilled in the art may be used to analyze the video information. MIPSS 104 may analyze the audio or closed-caption text information included in the multimedia document to identify sections of the information that include information relevant to user-specified topics of interest. For example, MIPSS 104 may generate a text transcript of the audio information and then analyze the text transcript to identify presence of words or phrases related to the user-specified topics of interest. Likewise, the CC text may also be analyzed. Other types of information (e.g., slides information, whiteboard information, etc.) included in the multimedia information stored by the multimedia document may also be analyzed. As previously stated, various analysis techniques known to those skilled in the art may be used to analyze the multimedia information stored by the multimedia document. MIPSS 104 may

30

then generate a printable representation for the multimedia document and annotate information in the printable representation that was deemed relevant to one or more user-specified topics of interest. The multimedia paper document generated by printing the annotated printable representation displays the annotations.

5           [159] Fig. 10 depicts a page 1000 whose contents have been annotated according to an embodiment of the present invention. Page 1000 depicted in Fig. 10 is merely illustrative of a multimedia paper document page and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives. As depicted in Fig. 10, words and phrases related to topics of interest are highlighted in text sections 1002 and 1004. For the embodiment depicted in Fig. 10 it is assumed that two topics of interest, namely "George Bush" and "Energy Crisis", have been specified. Keywords and phrases related to these topics of interest are highlighted. Different colors and styles (e.g., bolding, underlining, different font size, etc.) may be used to highlight words and phrases related to different topics. For example, as depicted in Fig. 10, a first color is used to highlight words and phrases related to the "George Bush" topic of interest and a second color is used to highlight words and phrases related to the "Energy Crisis" topic of interest.

1000  
1001  
1002  
1003  
1004  
1005  
1006  
1007  
1008  
1009  
1010  
1011  
1012  
1013  
1014  
1015  
1016  
1017  
1018  
1019  
1020  
1021  
1022  
1023  
1024  
1025  
1026  
1027  
1028  
1029  
1030  
1031  
1032  
1033  
1034  
1035  
1036  
1037  
1038  
1039  
1040  
1041  
1042  
1043  
1044  
1045  
1046  
1047  
1048  
1049  
1050  
1051  
1052  
1053  
1054  
1055  
1056  
1057  
1058  
1059  
1060  
1061  
1062  
1063  
1064  
1065  
1066  
1067  
1068  
1069  
1070  
1071  
1072  
1073  
1074  
1075  
1076  
1077  
1078  
1079  
1080  
1081  
1082  
1083  
1084  
1085  
1086  
1087  
1088  
1089  
1090  
1091  
1092  
1093  
1094  
1095  
1096  
1097  
1098  
1099  
1100  
1101  
1102  
1103  
1104  
1105  
1106  
1107  
1108  
1109  
1110  
1111  
1112  
1113  
1114  
1115  
1116  
1117  
1118  
1119  
1120  
1121  
1122  
1123  
1124  
1125  
1126  
1127  
1128  
1129  
1130  
1131  
1132  
1133  
1134  
1135  
1136  
1137  
1138  
1139  
1140  
1141  
1142  
1143  
1144  
1145  
1146  
1147  
1148  
1149  
1150  
1151  
1152  
1153  
1154  
1155  
1156  
1157  
1158  
1159  
1160  
1161  
1162  
1163  
1164  
1165  
1166  
1167  
1168  
1169  
1170  
1171  
1172  
1173  
1174  
1175  
1176  
1177  
1178  
1179  
1180  
1181  
1182  
1183  
1184  
1185  
1186  
1187  
1188  
1189  
1190  
1191  
1192  
1193  
1194  
1195  
1196  
1197  
1198  
1199  
1200  
1201  
1202  
1203  
1204  
1205  
1206  
1207  
1208  
1209  
1210  
1211  
1212  
1213  
1214  
1215  
1216  
1217  
1218  
1219  
1220  
1221  
1222  
1223  
1224  
1225  
1226  
1227  
1228  
1229  
1230  
1231  
1232  
1233  
1234  
1235  
1236  
1237  
1238  
1239  
1240  
1241  
1242  
1243  
1244  
1245  
1246  
1247  
1248  
1249  
1250  
1251  
1252  
1253  
1254  
1255  
1256  
1257  
1258  
1259  
1260  
1261  
1262  
1263  
1264  
1265  
1266  
1267  
1268  
1269  
1270  
1271  
1272  
1273  
1274  
1275  
1276  
1277  
1278  
1279  
1280  
1281  
1282  
1283  
1284  
1285  
1286  
1287  
1288  
1289  
1290  
1291  
1292  
1293  
1294  
1295  
1296  
1297  
1298  
1299  
1300  
1301  
1302  
1303  
1304  
1305  
1306  
1307  
1308  
1309  
1310  
1311  
1312  
1313  
1314  
1315  
1316  
1317  
1318  
1319  
1320  
1321  
1322  
1323  
1324  
1325  
1326  
1327  
1328  
1329  
1330  
1331  
1332  
1333  
1334  
1335  
1336  
1337  
1338  
1339  
1340  
1341  
1342  
1343  
1344  
1345  
1346  
1347  
1348  
1349  
1350  
1351  
1352  
1353  
1354  
1355  
1356  
1357  
1358  
1359  
1360  
1361  
1362  
1363  
1364  
1365  
1366  
1367  
1368  
1369  
1370  
1371  
1372  
1373  
1374  
1375  
1376  
1377  
1378  
1379  
1380  
1381  
1382  
1383  
1384  
1385  
1386  
1387  
1388  
1389  
1390  
1391  
1392  
1393  
1394  
1395  
1396  
1397  
1398  
1399  
1400  
1401  
1402  
1403  
1404  
1405  
1406  
1407  
1408  
1409  
1410  
1411  
1412  
1413  
1414  
1415  
1416  
1417  
1418  
1419  
1420  
1421  
1422  
1423  
1424  
1425  
1426  
1427  
1428  
1429  
1430  
1431  
1432  
1433  
1434  
1435  
1436  
1437  
1438  
1439  
1440  
1441  
1442  
1443  
1444  
1445  
1446  
1447  
1448  
1449  
1450  
1451  
1452  
1453  
1454  
1455  
1456  
1457  
1458  
1459  
1460  
1461  
1462  
1463  
1464  
1465  
1466  
1467  
1468  
1469  
1470  
1471  
1472  
1473  
1474  
1475  
1476  
1477  
1478  
1479  
1480  
1481  
1482  
1483  
1484  
1485  
1486  
1487  
1488  
1489  
1490  
1491  
1492  
1493  
1494  
1495  
1496  
1497  
1498  
1499  
1500  
1501  
1502  
1503  
1504  
1505  
1506  
1507  
1508  
1509  
1510  
1511  
1512  
1513  
1514  
1515  
1516  
1517  
1518  
1519  
1520  
1521  
1522  
1523  
1524  
1525  
1526  
1527  
1528  
1529  
1530  
1531  
1532  
1533  
1534  
1535  
1536  
1537  
1538  
1539  
1540  
1541  
1542  
1543  
1544  
1545  
1546  
1547  
1548  
1549  
1550  
1551  
1552  
1553  
1554  
1555  
1556  
1557  
1558  
1559  
1560  
1561  
1562  
1563  
1564  
1565  
1566  
1567  
1568  
1569  
1570  
1571  
1572  
1573  
1574  
1575  
1576  
1577  
1578  
1579  
1580  
1581  
1582  
1583  
1584  
1585  
1586  
1587  
1588  
1589  
1590  
1591  
1592  
1593  
1594  
1595  
1596  
1597  
1598  
1599  
1600  
1601  
1602  
1603  
1604  
1605  
1606  
1607  
1608  
1609  
1610  
1611  
1612  
1613  
1614  
1615  
1616  
1617  
1618  
1619  
1620  
1621  
1622  
1623  
1624  
1625  
1626  
1627  
1628  
1629  
1630  
1631  
1632  
1633  
1634  
1635  
1636  
1637  
1638  
1639  
1640  
1641  
1642  
1643  
1644  
1645  
1646  
1647  
1648  
1649  
1650  
1651  
1652  
1653  
1654  
1655  
1656  
1657  
1658  
1659  
1660  
1661  
1662  
1663  
1664  
1665  
1666  
1667  
1668  
1669  
1670  
1671  
1672  
1673  
1674  
1675  
1676  
1677  
1678  
1679  
1680  
1681  
1682  
1683  
1684  
1685  
1686  
1687  
1688  
1689  
1690  
1691  
1692  
1693  
1694  
1695  
1696  
1697  
1698  
1699  
1700  
1701  
1702  
1703  
1704  
1705  
1706  
1707  
1708  
1709  
1710  
1711  
1712  
1713  
1714  
1715  
1716  
1717  
1718  
1719  
1720  
1721  
1722  
1723  
1724  
1725  
1726  
1727  
1728  
1729  
1730  
1731  
1732  
1733  
1734  
1735  
1736  
1737  
1738  
1739  
1740  
1741  
1742  
1743  
1744  
1745  
1746  
1747  
1748  
1749  
1750  
1751  
1752  
1753  
1754  
1755  
1756  
1757  
1758  
1759  
1760  
1761  
1762  
1763  
1764  
1765  
1766  
1767  
1768  
1769  
1770  
1771  
1772  
1773  
1774  
1775  
1776  
1777  
1778  
1779  
1780  
1781  
1782  
1783  
1784  
1785  
1786  
1787  
1788  
1789  
1790  
1791  
1792  
1793  
1794  
1795  
1796  
1797  
1798  
1799  
1800  
1801  
1802  
1803  
1804  
1805  
1806  
1807  
1808  
1809  
1810  
1811  
1812  
1813  
1814  
1815  
1816  
1817  
1818  
1819  
1820  
1821  
1822  
1823  
1824  
1825  
1826  
1827  
1828  
1829  
1830  
1831  
1832  
1833  
1834  
1835  
1836  
1837  
1838  
1839  
1840  
1841  
1842  
1843  
1844  
1845  
1846  
1847  
1848  
1849  
1850  
1851  
1852  
1853  
1854  
1855  
1856  
1857  
1858  
1859  
1860  
1861  
1862  
1863  
1864  
1865  
1866  
1867  
1868  
1869  
1870  
1871  
1872  
1873  
1874  
1875  
1876  
1877  
1878  
1879  
1880  
1881  
1882  
1883  
1884  
1885  
1886  
1887  
1888  
1889  
1890  
1891  
1892  
1893  
1894  
1895  
1896  
1897  
1898  
1899  
1900  
1901  
1902  
1903  
1904  
1905  
1906  
1907  
1908  
1909  
1910  
1911  
1912  
1913  
1914  
1915  
1916  
1917  
1918  
1919  
1920  
1921  
1922  
1923  
1924  
1925  
1926  
1927  
1928  
1929  
1930  
1931  
1932  
1933  
1934  
1935  
1936  
1937  
1938  
1939  
1940  
1941  
1942  
1943  
1944  
1945  
1946  
1947  
1948  
1949  
1950  
1951  
1952  
1953  
1954  
1955  
1956  
1957  
1958  
1959  
1960  
1961  
1962  
1963  
1964  
1965  
1966  
1967  
1968  
1969  
1970  
1971  
1972  
1973  
1974  
1975  
1976  
1977  
1978  
1979  
1980  
1981  
1982  
1983  
1984  
1985  
1986  
1987  
1988  
1989  
1990  
1991  
1992  
1993  
1994  
1995  
1996  
1997  
1998  
1999  
2000  
2001  
2002  
2003  
2004  
2005  
2006  
2007  
2008  
2009  
2010  
2011  
2012  
2013  
2014  
2015  
2016  
2017  
2018  
2019  
2020  
2021  
2022  
2023  
2024  
2025  
2026  
2027  
2028  
2029  
2030  
2031  
2032  
2033  
2034  
2035  
2036  
2037  
2038  
2039  
2040  
2041  
2042  
2043  
2044  
2045  
2046  
2047  
2048  
2049  
2050  
2051  
2052  
2053  
2054  
2055  
2056  
2057  
2058  
2059  
2060  
2061  
2062  
2063  
2064  
2065  
2066  
2067  
2068  
2069  
2070  
2071  
2072  
2073  
2074  
2075  
2076  
2077  
2078  
2079  
2080  
2081  
2082  
2083  
2084  
2085  
2086  
2087  
2088  
2089  
2090  
2091  
2092  
2093  
2094  
2095  
2096  
2097  
2098  
2099  
2100  
2101  
2102  
2103  
2104  
2105  
2106  
2107  
2108  
2109  
2110  
2111  
2112  
2113  
2114  
2115  
2116  
2117  
2118  
2119  
2120  
2121  
2122  
2123  
2124  
2125  
2126  
2127  
2128  
2129  
2130  
2131  
2132  
2133  
2134  
2135  
2136  
2137  
2138  
2139  
2140  
2141  
2142  
2143  
2144  
2145  
2146  
2147  
2148  
2149  
2150  
2151  
2152  
2153  
2154  
2155  
2156  
2157  
2158  
2159  
2160  
2161  
2162  
2163  
2164  
2165  
2166  
2167  
2168  
2169  
2170  
2171  
2172  
2173  
2174  
2175  
2176  
2177  
2178  
2179  
2180  
2181  
2182  
2183  
2184  
2185  
2186  
2187  
2188  
2189  
2190  
2191  
2192  
2193  
2194  
2195  
2196  
2197  
2198  
2199  
2200  
2201  
2202  
2203  
2204  
2205  
2206  
2207  
2208  
2209  
2210  
2211  
2212  
2213  
2214  
2215  
2216  
2217  
2218  
2219  
2220  
2221  
2222  
2223  
2224  
2225  
2226  
2227  
2228  
2229  
2230  
2231  
2232  
2233  
2234  
2235  
2236  
2237  
2238  
2239  
2240  
2241  
2242  
2243  
2244  
2245  
2246  
2247  
2248  
2249  
2250  
2251  
2252  
2253  
2254  
2255  
2256  
2257  
2258  
2259  
2260  
2261  
2262  
2263  
2264  
2265  
2266  
2267  
2268  
2269  
2270  
2271  
2272  
2273  
2274  
2275  
2276  
2277  
2278  
2279  
2280  
2281  
2282  
2283  
2284  
2285  
2286  
2287  
2288  
2289  
2290  
2291  
2292  
2293  
2294  
2295  
2296  
2297  
2298  
2299  
2300  
2301  
2302  
2303  
2304  
2305  
2306  
2307  
2308  
2309  
2310  
2311  
2312  
2313  
2314  
2315  
2316  
2317  
2318  
2319  
2320  
2321  
2322  
2323  
2324  
2325  
2326  
2327  
2328  
2329  
2330  
2331  
2332  
2333  
2334  
2335  
2336  
2337  
2338  
2339  
2340  
2341  
2342  
2343  
2344  
2345  
2346  
2347  
2348  
2349  
2350  
2351  
2352  
2353  
2354  
2355  
2356  
2357  
2358  
2359  
2360  
2361  
2362  
2363  
2364  
2365  
2366  
2367  
2368  
2369  
2370  
2371  
2372  
2373  
2374  
2375  
2376  
2377  
2378  
2379  
2380  
2381  
2382  
2383  
2384  
2385  
2386  
2387  
2388  
2389  
2390  
2391  
2392  
2393  
2394  
2395  
2396  
2397  
2398  
2399  
2400  
2401  
2402  
2403  
2404  
2405  
2406  
2407  
2408  
2409  
2410  
2411  
2412  
2413  
2414  
2415  
2416  
2417  
2418  
2419  
2420  
2421  
2422  
2423  
2424  
2425  
2426  
2427  
2428  
2429  
2430  
2431  
2432  
2433  
2434  
2435  
2436  
2437  
2438  
2439  
2440  
2441  
2442  
2443  
2444  
2445  
2446  
2447  
2448  
2449  
2450  
2451  
2452  
2453  
2454  
2455  
2456  
2457  
2458  
2459  
2460  
2461  
2462  
2463  
2464  
2465  
2466  
2467  
2468  
2469  
2470  
2471  
2472  
2473  
2474  
2475  
2476  
2477  
2478  
2479  
2480  
2481  
2482  
2483  
2484  
2485  
2486  
2487  
2488  
2489  
2490  
2491  
2492  
2493  
2494  
2495  
2496  
2497  
2498  
2499  
2500  
2501  
2502  
2503  
2504  
2505  
2506  
2507  
2508  
2509  
2510  
2511  
2512  
2513  
2514  
2515  
2516  
2517  
2518  
2519  
2520  
2521  
2522  
2523  
2524  
2525  
2526  
2527  
2528  
2529  
2530  
2531  
2532  
2533  
2534  
2535  
2536  
2537  
2538  
2539  
2540  
2541  
2542  
2543  
2544  
2545  
2546  
2547  
2548  
2549  
2550  
2551  
2552  
2553  
2554  
2555  
2556  
2557  
2558  
2559  
2560  
2561  
2562  
2563  
2564  
2565  
2566  
2567  
2568  
2569  
2570  
2571  
2572  
2573  
2574  
2575  
2576  
2577  
2578  
2579  
2580  
2581  
2582  
2583  
2584  
2585  
2586  
2587  
2588  
2589  
2590  
2591  
2592  
2593  
2594  
2595  
2596  
2597  
2598  
2599  
2600  
2601  
2602  
2603  
2604  
2605  
2606  
2607  
2608  
2609  
2610  
2611  
2612  
2613  
2614  
2615  
2616  
2617  
2618  
2619  
2620  
2621  
2622  
2623  
2624  
2625  
2626  
2627  
2628  
2629  
2630  
2631  
2632  
2633  
2634  
2635  
2636  
2637  
2638  
2639  
2640  
2641  
2642  
2643  
2644  
2645  
2646  
2647  
2648  
2649  
2650  
2651  
2652  
2653  
2654  
2655  
2656  
2657  
2658  
2659  
2660  
2661  
2662  
2663  
2664  
2665  
2666  
2667  
2668  
2669  
2670  
2671  
2672  
2673  
2674  
2675  
2676  
2677  
2678  
2679  
2680  
2681  
2682  
2683  
2684  
2685  
2686  
2687  
2688  
2689  
2690  
2691  
2692  
2693  
2694  
2695  
2696  
2697  
2698  
2699  
2700  
2701  
2702  
2703  
2704  
2705  
2706  
2707  
2708  
2709  
2710  
2711  
2712  
2713  
2714  
2715  
2716  
2717  
2718  
2719  
2720  
2721  
2722  
2723  
2724  
2725  
2726  
2727  
2728  
2729  
2730  
2731  
2732  
2733  
2734  
2735  
2736  
2737  
2738  
2739  
2740  
2741  
2742  
2743  
2744  
2745  
2746  
2747  
2748  
2749  
2750  
2751  
2752  
2753  
2754  
2755  
2756  
2757  
2758  
2759  
2760  
2761  
2762  
2763  
2764  
2765  
2766  
2767  
2768  
2769  
2770  
2771  
2772  
2773  
2774  
2775  
2776  
2777  
2778  
2779  
2780  
2781  
2782  
2783  
2784  
2785  
2786  
2787  
2788  
2789  
2790  
2791  
2792  
2793  
2794  
2795  
2796  
2797  
2798  
2799  
2800  
2801  
2802  
2803  
2804  
2805  
2806  
2807  
2808  
2809  
2810  
2811  
2812  
2813  
2814  
2815  
2816  
2817  
2818  
2819  
2820  
2821  
2822  
2823  
2824  
2825  
2826  
2827  
2828  
2829  
2830  
2831  
2832  
2833  
2834  
2835  
2836  
2837  
2838  
2839  
2840  
2841  
2842  
2843  
2844  
2845  
2846  
2847  
2848  
2849  
2850  
2851  
2852  
2853  
2854  
2855  
2856  
2857  
2858  
2859  
2860  
2861  
2862  
2863  
2864  
2865  
2866  
2867  
2868  
2869  
2870  
2871  
2872  
2873  
2874  
2875  
2876  
2877  
2878  
2879  
2880  
2881  
2882  
2883  
2884  
2885  
2886  
2887  
2888  
2889  
2890  
2891  
2892  
2893  
2894  
2895  
2896  
2897  
2898  
2899  
2900  
2901  
2902  
2903  
2904  
2905  
2906  
2907  
2908  
2909  
2910  
2911  
2912  
2913  
2914  
2915  
2916  
2917  
2918  
2919  
2920  
2921  
2922  
2923  
2924  
2925  
2926  
2927  
2928  
2929  
2930  
2931  
2932  
2933  
2934  
2935  
2936  
2937  
2938  
2939  
2940  
2941  
2942  
2943  
2944  
2945  
2946  
2947  
2948  
2949  
2950  
2951  
2952  
2953  
2954  
2955  
2956  
2957  
2958  
2959  
2960  
2961  
2962  
2963  
2964  
2965  
2966  
2967  
2968  
2969  
2970  
2971  
2972  
2973  
2974  
2975  
2976  
2977  
2978  
2979  
2980  
2981  
2982  
2983  
2984  
2985  
2986  
2987  
2988  
2989  
2990  
2991  
2992  
2993  
2994  
2995  
2996  
2997  
2

for the various broadcasts, a user can easily determine the news broadcast that is most relevant to the user for any given user topic.

5 [162] According to an embodiment of the present invention, information stored in a user's profile and words or phrases related to user-specified topics of interest detected in the text section (CC text or transcript of audio information) may also be used to select keyframes from the video information that are relevant to the user-specified topics of interest. Since only a limited number of keyframes can be printed on the multimedia paper document due to limited amount of space available on a multimedia paper document page, selection of keyframes relevant to the user improves the readability of the document for the user.

[163] As described above, a user profile may be configured by a user and may store information identifying one or more topics of interest to the user. One or more words or phrases may be associated with each topic of interest such that presence of the words and phrases associated with a particular topic of interest indicates presence of information related to the particular topic. According to an embodiment of the present invention, probability values may be associated with each of the words or phrases indicating the likelihood of the topic of interest given the presence of the word or phrase. In order to facilitate selection of relevant keyframes, the user profile also stores information about features of the video information (or of the keyframes) that the user would like the present invention to search for in the video information when a word or phrase related to a topic is found in the text section.

25 [164] As previously described, several different features can be recognized from the video information. These features may include recognition of a human face, buildings, geographical locations, presence of a crowd, hand gestures, logos or signs, meetings, animals, text, and the like. Various algorithms known to those skilled in the art may be used to detect the features from the video information. For each of the features stated above, techniques that recognize video features may also be used to identify specific instances of a feature. For example, if a face is identified in a video frame, face recognition techniques may be applied to determine the identity of the face (i.e., a specific instance of a face). Likewise, if a logo was identified in a video frame, techniques may be applied to determine the name of the company corresponding to the logo. Similarly, if a building was identified in a video frame, techniques may be applied to determine if the building was a specific building such as the Empire State Building. Likewise, if an animal was identified in

30

a video frame, techniques may be applied to determine the type (e.g., horse, cat, dog, etc.) of animal.

[165] As part of a user's profile, the user may specify one or more video features (and specific instances of the features where appropriate) to be associated with one or more topics of interest. According to an embodiment of the present invention, the video features may be associated with keywords or phrases associated with user topics of interest. For each video feature, the user may also specify weight values for each topic indicating the relative importance of the video feature for that topic of interest.

[166] Fig. 11 depicts a user profile 1100 that may be configured by a user according to an embodiment of the present invention to facilitate selection of keyframes relevant to user-specified topics of interest. As depicted in Fig. 11, three topics of interest have been specified, namely, "Terrorism", "Company XYZ", and "Football". Keywords and phrases have been associated with each of the topics. In order to facilitate selection of keyframes relevant to the user topics of interest from the video information, the user has also specified video features to be searched for when keywords and phrases associated with the topics are located in the text (e.g., in the CC text or in the text transcript of audio information) of the multimedia information. Weights have been associated with the video features indicating the relative importance of the video features for each topic of interest. For example, for the topic "Terrorism", the face of Osama Bin Laden (weighted 0.7) is slightly more important than presence of text "Afghanistan" (weighted 0.6).

[167] Profile 1100 specifies the criteria for selecting keyframes relevant to the topics of interest given the presence of a keyword or phrase related to the topics of interest. For example, profile information 1100 specifies that the words "Osama" and "Afghanistan" are associated with the topic "Terrorism". If the word "Osama" is located in the text information of the multimedia information, then the video information (or video frames which have been extracted from the video information) temporally proximal to the occurrence of the word "Osama" in the text information are to be checked to determine if they include a face of Osama Bin Laden. Keyframes that contain Osama Bin Laden's face are deemed to be relevant (degree of relevance indicated by the weight value) to topic "Terrorism."

[168] Likewise, if the word "Afghanistan" is located in the text information of the multimedia information, then the video frames temporally proximal to the occurrence of the word "Afghanistan" in the text information are to be checked to determine if they contain text "Afghanistan". As previously described, OCR techniques may be used to extract



text from video keyframes. Keyframes that contain text "Afghanistan" are deemed to be relevant (degree of relevance indicated by the weight value) to topic "Terrorism."

[169] Further, for all (indicated by "\*") keywords and phrases (including "Osama" and "Afghanistan") associated with the topic "Terrorism," video frames temporally proximal to the occurrence of the words or phrases in the text information are to be checked to determine if they contain a building or (indicated by the Boolean connector OR) a crowd. Such keyframes are deemed to be relevant (degree of relevance indicated by the weight value) to topic "Terrorism." Accordingly, if the word "Osama" is located in the text information of the multimedia information, then the video frames temporally proximal to the occurrence of the word "Osama" in the text information would be first checked to determine if they include a face of Osama Bin Laden, and then checked to determine if they contain a building or a crowd.

[170] Likewise, profile information 1100 specifies that the word "Suzuki" is associated with the topic "Company XYZ" ("Suzuki" may be the name of the CEO of Company XYZ). If the word "Suzuki" is located in the text information of the multimedia information, then the video frames temporally proximal to the occurrence of the word "Suzuki" in the text information are to be checked to determine if they include a face of John Suzuki. Keyframes that contain John Suzuki's face are deemed to be relevant (degree of relevance indicated by the weight value) to topic "Company XYZ."

[171] Further, for all (indicated by "\*") keywords and phrases (including "Suzuki") associated with the topic "Company XYZ", video frames temporally proximal to the occurrence of the words or phrases are to be checked to determine if they contain a building and (indicated by the Boolean connector AND) further if they contain either a XYZ logo or (indicated by the Boolean connector OR) text "XYZ". Such keyframes are deemed to be relevant (degree of relevance indicated by the weight value) to topic "Company XYZ."

[172] Likewise, profile information 1100 specifies that the phrase "Buffalo Bills" is associated with the topic "Football". If the phrase "Buffalo Bills" is located in the text information of the multimedia information, then the video frames temporally proximal to the occurrence of the phrase are to be checked to determine if they include a face of Jim Kelly or the face of Marv Levy. Keyframes that contain either Jim Kelly's or Marv Levy's face are deemed to be relevant (degree of relevance indicated by the weight value) to topic "Football".

[173] Fig. 12 depicts modules that facilitate selection of keyframes relevant to topics of interest according to an embodiment of the present invention. The modules



depicted in Fig. 12 may be software modules, hardware modules, or combinations thereof. The modules depicted in Fig. 12 are merely illustrative of an embodiment of the present invention and are not meant to limit the scope of the present invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

5 [174] As depicted in Fig. 12, a video feature recognition module 1202 receives as input video frames corresponding to (or extracted from) the video information contained by a multimedia document. For each video frame, video feature recognition module 1202 determines if the video frame contains any video features such as a face, a building, a logo, text, etc. If a particular video feature is located, video feature recognition module 1202 assigns a probability value to the feature indicating the probability that the video frame contains the identified video feature. Video feature recognition module 1202 also determines a specific instance of the video feature and assigns a probability value to it. For example, for a particular video frame, video feature recognition module 1202 may determine that there is an 85% probability that the video frame contains a face and that there is a 90% probability that the face belongs to John Suzuki. For the same video frame, video feature recognition module 1202 may determine that there is only a 3% probability that the video frame contains a building, and only a 1% probability that the video frame contains a logo. The output of video feature recognition module 1202 is a ranked list of features and specific instances for each video frame. If no video feature is detected, a generic keyframe selection procedure may be applied. The procedure may calculate the probability that a frame is a potential keyframe. The video frames and their associated ranked list information is then forwarded to frame selector module 1204 for further processing.

10 25 [175] Profile matcher module 1206 receives as input user profile information and text information (e.g., CC text information or transcript of audio information) extracted from the multimedia document. Based upon the user profile information, profile matcher module 1206 searches the text information to locate words and phrases in the text information that are related to user-specified topics of interest. The words and phrases located in the text information are annotated by profile matcher module 1206. As described previously, various techniques may be used to annotate the text information. The text information along with the annotations is then forwarded to frames selector module 1204 and to printable representation generator module 1208 for further processing.

30 [176] As described above, frame selector module 1204 receives as input video frames and associated ranked list information from video feature recognition module

1202, the annotated text information from profile matcher module 1206, and user profile information. Based upon the user profile information, for each annotated word or phrase found in the text information, frame selector module 1204 determines video frames relevant to the topic with which the annotated word or phrase is associated.

5           [177] According to an embodiment of the present invention, for a particular annotated word, frame selector module 1204 identifies video frames temporally proximal to the occurrence of the particular annotated word in the text information. This may be performed by determining the time stamp associated with the annotated word, and identifying a set of video frames within a time window of  $\pm N$  seconds from the time stamp of the  
10 annotated word. The value of  $N$  is user configurable. For example, if  $N=5$  seconds, frame selector module 1204 identifies a set of video frames within a time window of  $\pm 5$  seconds from the time stamp of the annotated word.

          [178] Frame selector module 1204 then determines the topic of interest with which the annotated word is associated and the video features that are relevant to the particular annotated word or topic of interest (as specified in the user profile). Each video frame in the set of video frames within the  $\pm N$  seconds time window is then searched to determine if it contains one or more video features specified in the user profile for the topic or word. A relevancy score is calculated for each video frame in the set of video frames.

          [179] According to an embodiment of the present invention, in order to calculate a relevancy score for a video frame, frame selector module 1204 multiplies the weight assigned to the video feature in the user profile by the probability value assigned to the video frame by video feature recognition module 1202 indicating that the frame contains the particular video feature. The other probability values assigned to the video frame by video feature recognition module 1202 may be multiplied by a constant ( $K$ ) that is less than  
25 the weights in the profile information. This ensures that the simultaneous detection of a keyword and a relevant video frame will provide a higher rank for that video frame than if the keyword was not detected. After each video frame in the set of video frames has been assigned a relevancy value, the video frames are ranked based upon their relevancy values. Accordingly, for each annotated word or phrase in the text information, frame selector  
30 module generates a ranked list of video frames. The ranked list of keyframes for each annotated word or phrase is then forwarded to printable representation generator module 1208 for further processing.

          [180] Printable representation generator module 1208 receives annotated text information from profile matcher module 1206 and ranked lists of keyframes for the

annotations from frame selector module 1204. Printable representation generator module also receives as input other types of information included in the multimedia information stored by the multimedia document and layout and format information. Based upon the various inputs, printable representation generator module 1208 generates a printable representation for the multimedia document.

[181] According to an embodiment of the present invention, as part of processing performed to generate the printable representation for the multimedia document, printable representation generator module 1208 determines which keyframes to be included in the printable representation for the multimedia document for each segment of the multimedia document based upon the layout and format information and the ranked listing received from frame selector module 1204. For example, let's assume that the layout information specifies that four keyframes are to be printed for each segment. In this scenario, if the text corresponding to a particular segment contains only one annotated word, then the top four most relevant keyframes from the ranked list of keyframes associated with the annotated word and received from frame selector module 1204 are selected to be included in the printable representation for the segment. If a particular segment contains four different annotated words, then printable representation generator module 1208 may select only the most relevant keyframe from the ranked lists of keyframes associated with each of the four annotated words for inclusion in the printable representation (making a total of 4 keyframes). Accordingly, printable representation generator module 1208 determines the keyframes to be included in the printable representation for each segment of the multimedia document using the ranked list of keyframes received from frame selector module 1208.

#### [182] USING A MULTIMEDIA PAPER DOCUMENT TO RETRIEVE MULTIMEDIA INFORMATION

[183] The present invention provides techniques that allow a user to access or retrieve multimedia information in digital form using the multimedia paper document generated for a multimedia document. The multimedia paper document may thus be used as an indexing and retrieval tool for retrieving multimedia information that may be stored in the multimedia document. For example, a user may use a multimedia paper document generated for a video recording to access or retrieve portions of the video recording.

[184] Fig. 13A is a simplified high-level flowchart 1300 depicting a method of retrieving multimedia information using a multimedia paper document according to an embodiment of the present invention. Flowchart 1300 depicted in Fig. 13 is merely

illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[185] As depicted in Fig. 13, a user may initiate the method by selecting one or more segments from a multimedia paper document corresponding to multimedia information that the user wishes to access or retrieve (step 1302). The segments may be selected by selecting user-selectable identifiers (e.g., user-selectable identifiers 726 depicted in Fig. 7A) associated with the segments using a selection device. The user-selectable identifiers corresponding to the segments may be selected from one or more pages of a multimedia paper document. Further, the user-selectable identifiers may be selected from one or more multimedia paper documents. Several different techniques may be provided by a multimedia paper document to enable the user to select one or more segments.

[186] According to an embodiment of the present invention, barcode technology is used to facilitate selection of segments. In this embodiment, the user-selectable identifiers are printed in the multimedia paper document in the form of barcodes. Each barcode corresponds to a particular segment of the multimedia document. For example, as depicted in Fig. 7A, according to an embodiment of the present invention, a barcode 726 is printed for each segment printed on a page of the multimedia paper document. For example, barcode 726-1 printed on page 700 corresponds to segment 1, barcode 726-2 corresponds to the segment 2, barcode 726-3 corresponds to the segment 3, and so on. A user can select a particular segment by scanning the barcode corresponding to that segment. A selection device such as a barcode scanner or any other device that is capable of scanning barcodes may be used to scan the barcodes. The user may scan one or more barcodes from one or more pages of one or more multimedia paper documents.

[187] It should be apparent that various other techniques, besides barcodes, may be used to facilitate selection of segments corresponding to multimedia information that the user wishes to access or retrieve. According to an embodiment of the present invention, the user-selectable identifiers may be implemented as watermarks printed on pages of the multimedia paper document may also be used as techniques for selecting segments. In this embodiment, a user may select one or more watermarks corresponding to segments of interest to the user using a selection device that is capable of reading or detecting the watermarks.

[188] According to another embodiment of the present invention, the user-selectable identifiers may be implemented as text string identifiers printed on pages of the multimedia paper document. In this embodiment, a user may select a particular segment by

keying in or entering the text string identifier corresponding to the particular segment into a selection device such as a telephone, a DVR, etc.

[189] Various other techniques (e.g., Xerox glyphs embedded in keyframes, etc.) known to those skilled in the art may also be used to facilitate selection of segments.

5 Generally, in order to maintain the readability of the multimedia paper document, techniques that are less obtrusive, and those that do not take up too much space on the page, and which are somewhat aesthetically pleasing may be used.

10 [190] After the user has selected one or more segments, the user may select preferences for playing back the multimedia information corresponding to the segments selected in step 1302 (step 1304). According to an embodiment of the present invention, the user may specify preferences by selecting one or more controls from controls section 710. As with selection of segments, various different techniques may be used to facilitate selection of controls. For example, according to an embodiment of the present invention, a particular control may be selected by scanning a barcode corresponding to the control. For example, the user may specify that the multimedia information is to be played back in "Enhanced Mode" by selecting barcode 724-4 depicted in Fig. 7A. The user may specify that the playback is to show CC text by selecting barcode 724-5 corresponding to control "Show Closed-caption". The user may specify that time is to be displayed during the playback by selecting barcode 724-6 corresponding to control "Show Time". The user in step 1304 may also select various other preferences.

[191] According to an embodiment of the present invention, as part of step 1304, the user may also specify an output device to be used for playing back the multimedia information corresponding to the segments selected in step 1302. According to an embodiment of the present invention, one or more devices that may be located at different geographical locations may be selected for playback. For example, the selected output device may be the user's PDA, a computer in the user's office, a television at the user's home, a specific kiosk, and the like.

25 [192] In alternative embodiments of the present invention, the user preferences and the output device may be pre-configured. For example, this information may be stored in a user profile. Alternatively, the preferences and the output device may default to some preset values. In such a scenario, step 1304 may not be performed.

30 [193] The user may then request playback of the multimedia information corresponding to the segments selected in step 1302 (step 1306). According to an embodiment of the present invention, the user may request playback by selecting a barcode

such as barcode 724-1 corresponding to the "Play" control. According to an embodiment of the present invention, upon selecting the "Play" control, a signal is transmitted from the selection device (e.g., a barcode scanner) used by the user to select the segments and the preferences to a server that is capable of retrieving multimedia information corresponding to the user-selected segments. The server may be MIPSS 104 or any other server. The signal communicated to the server from the selection device may identify the segments selected by the user in step 1302, the multimedia paper documents from which the segments are to be selected, information related to preferences and/or one or more output devices selected by the user in step 1304, and other like information to facilitate retrieval of the requested multimedia information.

[194] Various techniques and communication links may be used to communicate the signal from the selection device used by the user to the server. For example, if the selection device is a barcode scanner, a communication link may be established between the barcode scanner and the server and the signal information may be communication to the server via the communication link. Different types of communication links may be used including hardwire links, optical links, satellite or other wireless communications links, wave propagation links, or any other mechanisms for communication of information. Various communication protocols may be used to facilitate communication of the signal via the communication links. These communication protocols may include TCP/IP, HTTP protocols, extensible markup language (XML), wireless application protocol (WAP), protocols under development by industry standard organizations, vendor-specific protocols, customized protocols, and others.

[195] In other embodiments, a telephone may be used as a selection device. For example, a user may use a telephone to establish a communication link with the server. The user may then communication the signal information to server using the telephone. For example, the user may key in user-selectable identifiers (e.g., text string identifiers) corresponding to the selected segments and controls using the telephone. Various other techniques may also be used to communicate the information to the server.

[196] The server receiving the signal from the selection device may then retrieve multimedia information corresponding to the user-selected segments (step 1308). According to an embodiment of the present invention, the server determines the user-selectable identifiers selected by the user. The server then determines segments of the multimedia document corresponding to the user-selectable identifiers selected by the user. The server then retrieves multimedia information corresponding to the selected segments.

[197] The multimedia information may be retrieved from a single multimedia document or from multiple multimedia documents. For example, if the user selected user-selectable identifiers from multiple multimedia documents, then the server retrieves multimedia information corresponding to selected segments from multiple multimedia documents.

[198] The multimedia information retrieved by the server is then communicated to the one or more output devices selected for playback (step 1310). The multimedia information is then played on the one or more output devices selected for playback (step 1312). The user may control playback of the multimedia information by selecting one more controls from control area 710 depicted in Fig. 7A. For example, the user may stop playback of the multimedia information by selecting barcode 724-1 corresponding to the "Stop" control. A user may fast-forward 10 seconds of the multimedia information by selecting barcode 724-2. A user may rewind 10 seconds of the multimedia information by selecting barcode 724-3. Various other controls not shown in Fig. 7A may also be provided in alternative embodiments of the present invention to control playback of the multimedia information.

[199] According to an alternative embodiment of the present invention, a user may use the multimedia paper document to start playing multimedia information from a user-selected time point in the multimedia document. In this embodiment, the user-selectable identifiers (e.g., barcodes 726 depicted in Fig. 7A) printed in a multimedia paper document represent particular time points in the multimedia document. According to an embodiment of the present invention, the barcodes may correspond to the same time points as the identifiers (e.g., identifiers 712 depicted in Fig. 7A) printed on a page of the multimedia paper document.

[200] Fig. 13B is a simplified high-level flowchart 1350 depicting a method of retrieving multimedia information from a particular time point using a multimedia paper document according to an embodiment of the present invention. Flowchart 1350 depicted in Fig. 13B is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[201] As depicted in Fig. 13, a user may initiate the method by selecting a user-selectable identifier printed on a multimedia paper document corresponding to a time point in the multimedia document from where the user wishes to retrieve multimedia information (step 1352). As described above, several different techniques (e.g., barcodes,



watermarks, glyphs, text strings, etc.) may be provided by the multimedia paper document to enable the user to the user-selectable identifier.

[202] After selecting a user-selectable identifier, the user may select preferences for playing back the multimedia information (step 1304). As described above with respect to Fig. 13A, the user may select a mode for playing back the multimedia information, select one or more output devices for playing back the multimedia information, and the like. Step 1354 may not be performed if the user preferences are pre-configured.

[203] The user may then request playback of the multimedia information (step 1356). According to an embodiment of the present invention, upon selecting the "Play" control, a signal is transmitted from the selection device (e.g., a barcode scanner) used by the user to a server that is capable of retrieving multimedia information from the multimedia document. The server may be MIPSS 104 or any other server. The signal communicated to the server from the selection device may identify the user-selectable identifier selected by the user in step 1352, the multimedia paper document from which the user-selectable identifier was selected, information related to preferences and/or one or more output devices selected by the user in step 1354, and other like information to facilitate retrieval of the requested multimedia information.

[204] The server receiving the signal from the selection device then retrieves multimedia information from the time point corresponding to the user-selectable identifier selected by the user in step 1352 (step 1358). According to an embodiment of the present invention, the server determines a time point in the multimedia document corresponding to the user-selectable identifier selected by the user and retrieves multimedia information from the time point onwards. view

[205] The multimedia information retrieved by the server in step 1358 is then communicated to the one or more output devices selected for playback (step 1360). The multimedia information is then played back on the one or more output devices selected for playback (step 1362). The user may control playback of the multimedia information by selecting one more controls from control area 710 depicted in Fig. 7A. For example, the user may stop playback of the multimedia information by selecting barcode 724-1 corresponding to the "Stop" control. A user may fast-forward 10 seconds of the multimedia information by selecting barcode 724-2. A user may rewind 10 seconds of the multimedia information by selecting barcode 724-3. Various other controls not shown in Fig. 7A may also be provided in alternative embodiments of the present invention to control playback of the multimedia information.



[206] Accordingly, as described above, the multimedia paper document provides a simple and easy-to-use mechanism for retrieving multimedia information. The convenience afforded by the multimedia paper document in retrieving multimedia information might be illustrated by the following example. Let's assume that a user has requested that the television program "Bloomberg" be recorded between the hours of 9-11AM during which important financial news is broadcast. Various different devices may be used to record the news broadcast including a video recorder, a digital video recorder (DVR) (e.g., a TIVO box), and the like. The user may then generate a multimedia paper document for the recorded news broadcast.

[207] Let's further assume that the user has 15 minutes before a power lunch with a client to digest the two-hour Bloomberg TV program to find out if any relevant information was mentioned regarding the client's company or their main competitor. With the paper-based version of the broadcast (i.e., the multimedia paper document), the user can quickly skim the multimedia paper document for relevant information. When the user finds one or more segments in the multimedia paper document of interest, the user can use a barcode scanner to scan the barcodes corresponding to segments in the multimedia paper document. The user may also scan a control barcode instructing the recorder to launch the video corresponding to the selected segments on a television in the user's office. This sends a message to the recorder to launch the video corresponding to the selected segments on the television in the user's office. If the user has selected multiple segments, multimedia information corresponding to the selected segments will be played on the user's television, skipping the segments or sections that are not selected by the user. In this manner, the user can quickly navigate two-hours of a television broadcast in 15 minutes watching only those portions of the broadcast that are of interest to the user, and be ready for the client meeting in time.

[208] In the above scenario the user could have selected segments from multiple multimedia paper documents generated for a plurality of news broadcasts from news agencies such as CNBC, CNN/fn, MSNBC, and the like. The user could then skim the multimedia paper documents to locate news related to the client or the client's competitors from the various broadcasts. This is equivalent to watching several hours of video in a short time—something that is very difficult to achieve if the user only has access to a video player. The user may then select segments of interest from the multiple multimedia papers and watch video corresponding to the selected segments.

[209] In the above scenario, the present invention automatically records a desired broadcast program based a user's profile and produces a multimedia paper document that acts both as a familiar skimming tool and a retrieval device for viewing desired portions of the video. In the above-described scenario, the interface is not on the user's personal computer--instead, the interface is in the user's hands in the form of paper. In some cases, this is a more desired environment since most individuals are familiar with and more comfortable with reading and using paper. The paper-based interface thus provides a unique mechanism for indexing or referring back to the digitized multimedia information stored by the multimedia document. The indexing technique provided by the present invention may then be used by a user to retrieve the multimedia information in digital format. The multimedia paper document provides a portable means for random access to the multimedia information, a task that traditionally required tedious searching of the multimedia information.

[210] GENERATING A SINGLE PRINTABLE REPRESENTATION FOR A PLURALITY OF MULTIMEDIA DOCUMENTS

[211] The present invention provides techniques for generating a single printable representation that includes multimedia information extracted from a plurality of different multimedia documents or multimedia sources. According to an embodiment of the present invention, the single printable representation includes multimedia information selected from the plurality of multimedia documents based upon selection criteria. A user may specify the selection criteria. The selection criteria may be based upon any attributes of the multimedia documents or their contents, or upon user-specified topics of interest, and the like. For example, the selection criteria may specify a particular subject (e.g., information related to the Taliban in Afghanistan, or abortion related information, etc.), a specified story line, and the like.

[212] For example, a user may specify that a single printable representation (or a single multimedia paper document) be generated consolidating stories and articles related to "Middle East Terrorism" from a plurality of news broadcast recordings. In response, the present invention generates a single printable representation that includes multimedia information from the plurality of news broadcast recordings related to "Middle East Terrorism." The single consolidated printable representation may then be printed to generate a single consolidated multimedia paper document that contains information related to "Middle East Terrorism" from multiple multimedia documents.

[213] According to another example, topics of interest to the user (which may be stored in a user profile) may be specified as the selection criteria. Based upon such a selection criteria, MIPSS 104 may generate a single printable representation that includes multimedia information from the plurality of news broadcast recordings related to the user-specified topics of interest. The single consolidated printable representation may then be printed to generate a single consolidated multimedia paper document that contains information related to "Middle East Terrorism" extracted from multiple multimedia documents. In this manner, multimedia information from various multimedia sources or documents related to user-specified topics of interest may be consolidated into a single printable representation that may then be printed to generate a multimedia paper document. The multimedia paper document generated in this manner is a valuable tool that enables the user to read and comprehend related information from multiple sources in a timely and efficient manner.

[214] Fig. 14 is a simplified high-level flowchart 1400 depicting a method of generating a single printable representation according to an embodiment of the present invention that includes multimedia information selected from a plurality of multimedia documents by analyzing the printable representations of the plurality of multimedia documents. The method depicted in Fig. 14 may be used to generate a single multimedia paper document including multimedia information selected from a plurality of multimedia documents. The processing depicted in Fig. 14 may be performed by software modules executing on MIPSS 104, by hardware modules coupled to MIPSS 104, or a combination thereof. In alternative embodiments of the present invention, the processing may be distributed among the various systems depicted in Fig. 1. The processing depicted in Fig. 14 is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[215] The method is initiated by determining the selection criteria (or criterion) to be used for selecting the multimedia information to be included in the single printable representation and by determining the plurality of multimedia documents (or multimedia sources) from which the multimedia information is to be selected or extracted (step 1402). MIPSS 104 then generates a printable representation for each multimedia document determined in step 1402 if a printable representation does not already exist for the multimedia document (step 1404). The printable representations for the multimedia documents may be generated according to the methods depicted in Figs. 4 and 6.

5 [216] For each multimedia document identified in step 1402, MIPSS 104 searches the pages from the printable representation of the multimedia document to identify a set of pages that comprise information that satisfies the selection criteria determined in step 1402 (step 1406). MIPSS 104 then generates a single consolidated printable representation that includes the pages determined in step 1406 (step 1408). The single printable representation generated in step 1408 may then be printed on a paper medium to generate a consolidated multimedia paper document (step 1410). The multimedia paper document generated in step 1410 comprises information selected from the plurality of multimedia documents based upon the selection criteria. For each page of the multimedia paper document generated in step 1410, information printed information that satisfies the selection criteria may be annotated.

10 [217] As described above, the printable representations of the multimedia documents are analyzed to identify portions of multimedia information from the various multimedia documents to be included in the consolidated printable representation. According to alternative embodiments of the present invention, the multimedia information stored by the multimedia documents may be analyzed to identify portions of the multimedia information that satisfy the selection criteria. A consolidated printable representation may then be generated to include portions of multimedia information from the various multimedia documents that satisfy the selection criteria. The consolidated printable representation may then be printed on a paper medium to generate a consolidated or "customized" multimedia paper document.

25 [218] Fig. 15 is a simplified high-level flowchart 1500 depicting another method of generating a single printable representation that includes information extracted from a plurality of multimedia documents by analyzing the multimedia information stored by the plurality of multimedia documents according to an embodiment of the present invention. The method depicted in Fig. 15 may be used to generate a single multimedia paper document including multimedia information extracted from a plurality of multimedia documents. The processing depicted in Fig. 15 may be performed by software modules executing on MIPSS 104, by hardware modules coupled to MIPSS 104, or a combination thereof. In alternative  
30 embodiments of the present invention, the processing may be distributed among the various systems depicted in Fig. 1. The processing depicted in Fig. 15 is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[219] The method is initiated by determining the selection criteria (or criterion) to be used for selecting the multimedia information to be included in the single printable representation and by determining the plurality of multimedia documents (or multimedia sources) from which the multimedia information is to be selected (step 1502).

5 For each multimedia document determined in step 1502, MIPSS 104 divides the multimedia information contained by the multimedia document into segments of a particular time length (step 1504). The process of dividing a multimedia document into segments has been described earlier with respect to Figs. 6.

10 [220] For each multimedia document identified in step 1502, MIPSS 104 then determines those segments or portions of the multimedia document that comprise information that satisfies the selection criteria identified in step 1502 (step 1506). MIPSS 104 then generates a single consolidated printable representation based upon the segments determined in step 1506 (step 1508). The single consolidated printable representation includes segments determined in step 1506. The single printable representation generated in step 1508 may then be printed on a paper medium to generate a consolidated multimedia paper document (step 1510). The multimedia paper document generated in step 1510 comprises information selected from the plurality of multimedia documents based upon the selection criteria. The multimedia paper document generated in step 1510 may comprise annotations identifying printed information that satisfies the selection criteria.

[221] A multimedia paper document generated according to the flowcharts depicted in Figs. 14 and 15 may then be used as any other multimedia paper document. For example, a user may select one or more user-selectable identifiers from the consolidated multimedia paper document (as described above) and retrieve multimedia information corresponding to segments associated with the user-selectable identifiers selected by the user.

25 [222] Figs. 16A, 16B, 16C, and 16D depict pages of a multimedia paper document generated according to an embodiment of the present invention using the method depicted in Fig. 14. The pages have been selected from a plurality of multimedia documents because they contain information related to the topic of interest "Middle East Terrorism" that was specified as the selection criteria. The pages have been selected from printable  
30 representations generated for a plurality of multimedia documents. For example, pages 1600 and 1602 depicted in Figs. 16A and 16B have been selected from a printable representation generated for a "CNN News Site (Channel 203)" recording that was recorded on May 30, 2001 starting at 12:59PM and is of length 56:40 minutes, page 1606 depicted in Fig. 16C has been selected from a printable representation generated for a "Newshour (PBS, Channel

233)" recording that was recorded on June 5, 2001 starting at 6:01PM and is of length 54:49 minutes, page 1604 depicted in Fig. 16D has been selected from a printable representation generated for a "Hardball (CNBC, Channel 356)" recording that was recorded on September 14, 2001 starting at 5:00PM and is of length 59:59 minutes. For each page, information related to "Middle East Terrorism" has been annotated. This enhances the readability of the multimedia paper document. Accordingly, information related to "Middle East Terrorism" from a plurality of multimedia documents is consolidated into one document.

[223] As described above, a user may generate a "customized" multimedia paper document by specifying appropriate selection criteria. In this manner, the user can quickly extract relevant information from multiple hours of multimedia broadcasts by simply reading the customized multimedia paper document. The present invention thus reduces the time spent by the user in locating and retrieving relevant information from multiple multimedia information sources or recordings.

#### [224] COVERSHEETS

[225] According to an embodiment of the present invention, the present invention also provides techniques for generating a coversheet for a multimedia paper document. The coversheet may provide a summary of the contents printed in the multimedia paper document.

[226] Fig. 17 depicts a coversheet 1700 generated for a multimedia paper document according to an embodiment of the present invention. Coversheet 1700 depicted in Fig. 17 is merely illustrative of a coversheet according to an embodiment of the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[227] As shown in Fig. 1700, coversheet 1700 comprises thumbnail images of individual pages included in the multimedia paper document. As shown, eight thumbnail images 1704 are printed on coversheet 1700 thereby indicating that the multimedia paper document comprises eight pages. A title section 1702 is also printed on coversheet 1700. Title section 1702 displays the source of the multimedia information (which may correspond to the filename of the multimedia document), the time and date when the multimedia information was recorded, and the total time of the recording.

[228] A barcode 1706 is associated with each thumbnail image. A user may use barcodes 1706 to access or retrieve multimedia information printed on the pages in digital form. If the user wishes to access the multimedia information corresponding to information

printed on a particular multimedia paper document page, the user may scan the barcode corresponding to that particular page and then access or retrieve the information in electronic form using an output device specified by the user. In this embodiment, selecting a barcode associated with a particular page is equivalent to selecting all the segments printed on that particular page. For example, if the user wishes to access multimedia information corresponding to the information printed on page 6 of the multimedia paper document, the user may scan barcode 1706-6 and then access the information (as previously described) using an output device. The user may select one or more barcodes from coversheet 1700.

[229] According to another embodiment of the present invention, a barcode 1706 associated with a particular page is the same as the barcode corresponding to the first segment printed on the particular page. In this embodiment, the user may scan a barcode for a particular page and retrieve multimedia information starting from the top of the particular page.

[230] Fig. 18 depicts a coversheet 1800 generated for a multimedia paper document according to another embodiment of the present invention. In addition to the features included in coversheet 1700 depicted in Fig. 17, coversheet 1800 displays a list of sentences 1804 for each thumbnail image 1802. According to an embodiment of the present invention, the sentences displayed for a particular thumbnail image summarize the contents of the page corresponding to the particular thumbnail image. Several different techniques may be used to select the sentences for a particular thumbnail image. According to an embodiment of the present invention, the first text sentence of each segment printed on the page corresponding to the thumbnail image may be printed in 1804. According to another embodiment of the present invention, segments that contains CC text with story-line separators (e.g., ">>>"), the first sentence of each story printed on the page corresponding to the thumbnail image may be printed in 1804. Other techniques known to those skilled in the art may also be used to determine the text to be printed in area 1804 of coversheet 1800.

[231] It should be apparent that coversheet 1800 depicted in Fig. 18 is merely illustrative of a coversheet according to an embodiment of the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[232] Fig. 19 depicts a coversheet 1900 generated according to another embodiment of the present invention for a multimedia paper document that has been annotated based upon user-specified topics of interest. A title section 1902 is printed on coversheet 1900 displaying the source of the multimedia information (which may correspond



to the filename of the multimedia document), the time and date when the multimedia information was recorded, and the total time of the recording. Topics of interest 1904 to which the multimedia paper document is relevant are also displayed. For each topic of interest, the degree of relevancy of the multimedia paper document to the topic of interest is also displayed. In the embodiment depicted in Fig. 1900, the degree or relevancy is denoted by a percentage value 1906.

[233] Coversheet 1900 displays a thumbnail image 1908 of each page included in the multimedia paper document. For pages that comprise information related to user-specified topics, the thumbnail images corresponding to those pages display the annotated words or phrases related to user-specified topics of interest. For a particular page comprising information related to one or more user-specified topics of interest, the number of hits 1910 related to the topics of interest found on the particular page are also displayed next to the thumbnail image of the page. Different colors and styles may be used to highlight words and phrases in the thumbnails related to different topics. The hits for a particular topic of interest may also be displayed using a color that is associated with the topic of interest and used to highlight words and phrases related to the topic of interest. This allows the user of the multimedia paper document to easily identify pages of the multimedia paper document that include information related to user-specified topics of interest.

[234] It should be apparent that coversheet 1900 depicted in Fig. 19 is merely illustrative of a coversheet according to an embodiment of the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[235] Fig. 20 depicts a coversheet 2000 generated according to an embodiment of the present invention for a multimedia paper document that includes pages selected from multiple multimedia paper documents based upon selection criteria. For example, the multimedia paper document may be generated according to flowchart 1400 depicted in Fig. 14. Coversheet 2000 depicted in Fig. 20 has been generated for a multimedia paper document that includes pages 1600, 1602, 1604, and 1606 depicted in Figs. 16A, 16B, 16C, and 16D, respectively. It should be apparent that coversheet 2000 depicted in Fig. 20 is merely illustrative of a coversheet according to an embodiment of the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[236] As depicted in Fig. 20, the selection criteria 2002 used for generating the multimedia paper document is printed on page 2000. Coversheet 2000 displays a



thumbnail image 2004 of each page included in the multimedia paper document. For pages that comprise information related to the search criteria, the thumbnail images corresponding to those pages displaying the information with annotations. The number of hits 2006 for pages are also displayed. A barcode 2008 associated with each page is also displayed.

- 5 Coversheet 2000 also displays a date range 2010 that may be selected by the user as part of the selection criteria. For example, multimedia paper document comprises information in the date range May 1, 2001 to September 20, 2001.

[237] Fig. 21 depicts another coversheet 2100 generated according to an embodiment of the present invention for a multimedia paper document that includes pages selected from multiple multimedia paper documents based upon some selection criteria. Coversheet 2100 depicted in Fig. 21 has been generated for a multimedia paper document that includes pages 1600, 1602, 1604, and 1606 depicted in Figs. 16A, 16B, 16C, and 16D, respectively. It should be apparent that coversheet 2000 depicted in Fig. 20 is merely illustrative of a coversheet according to an embodiment of the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[238] Coversheet 2100 shows more information than coversheet 2000 depicted in Fig. 2000. For each occurrence of words or phrases related to the selection criteria (e.g., text related to "Middle East Terrorism"), the line 2102 (or a user-configurable number of words surrounding the relevant word/phrase) comprising the relevant text or phrase (which is annotated) is displayed along with the time 2104 when the word/phrase occurred in the recording and the page 2106 of the multimedia paper document on which the line is printed.

- [239] A barcode 2108 is also displayed for each line. According to an embodiment of the present invention, barcode 2108 corresponds to the barcode for the page on which the line occurs. According to alternative embodiments of the present invention, the barcode 2108 associated with a line may correspond to the barcode of the segment that contains the displayed line. Alternatively, barcode 2108 may correspond to a location within the multimedia information when the relevant text/phrase occurs. Accordingly, barcode 2108 enables the user to access or retrieve multimedia information from a specific point in time.

[240] A set of keyframes 2110 is also displayed for each line. According to an embodiment of the present invention, the keyframes that are most representative of the word/phrase or are relevant to the selection criteria may be displayed. Techniques for

selecting keyframes relevant to selection criteria such as a user-specified topic of interest have been described above.

[241] Fig. 22 depicts a coversheet 2200 generated according to an embodiment of the present invention for a multimedia paper document that has been generated for a recorded meeting. As shown in Fig. 22, coversheet 2200 comprises thumbnail images of individual pages included in the multimedia paper document. As shown, six thumbnail images 2202 are printed on coversheet 2200 thereby indicating that the multimedia paper document comprises eight pages. A title section 2204 is also printed on coversheet 2200 and displays information identifying the meeting for which the multimedia paper document was generated, the time and date when the meeting was recorded, and the total time of the recording. Slides 2206 and whiteboard images 2208 are also printed next to thumbnail images corresponding to pages that comprise the slides or whiteboard images.

[242] It should be apparent that coversheets 1700, 1800, 1900, 2000, 2100, and 2200 depicted in Figs. 17, 18, 19, 20, 21, and 22, respectively, are merely illustrative of specific embodiments of the present invention and do not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives. The coversheet generated according to the teachings of the present invention thus provide a simple and convenient way for the reader of the multimedia paper document to get an overview of the contents of the multimedia paper document.

[243] Although specific embodiments of the invention have been described, various modifications, alterations, alternative constructions, and equivalents are also encompassed within the scope of the invention. The described invention is not restricted to operation within certain specific data processing environments, but is free to operate within a plurality of data processing environments. Additionally, although the present invention has been described using a particular series of transactions and steps, it should be apparent to those skilled in the art that the scope of the present invention is not limited to the described series of transactions and steps.

[244] Further, while the present invention has been described using a particular combination of hardware and software, it should be recognized that other combinations of hardware and software are also within the scope of the present invention. The present invention may be implemented only in hardware, or only in software, or using combinations thereof.

